Neural Network Model for Construction Cost Prediction of Apartment Projects in Vietnam

Van Truong Luu* Soo-Yong Kim**

Abstract

Accurate construction cost estimation in the initial stage of building project plays a key role for project success and for mitigation of disputes. Total construction cost(TCC) estimation of apartment projects in Vietnam has become more important because those projects increasingly rise in quantity with the urbanization and population growth. This paper presents the application of artificial neural networks(ANNs) in estimating TCC of apartment projects. Ninety-one questionnaires were collected to identify input variables. Fourteen data sets of completed apartment projects were obtained and processed for training and generalizing the neural network(NN). MATLAB software was used to train the NN. A program was constructed using Visual C++ in order to apply the neural network to realistic projects. The results suggest that this model is reasonable in predicting TCCs for apartment projects and reinforce the reliability of using neural networks to cost models. Although the proposed model is not validated in a rigorous way, the ANN-based model may be useful for both practitioners and researchers. It facilitates systematic predictions in early phases of construction projects. Practitioners are more proactive in estimating construction costs and making consistent decisions in initial phases of apartment projects. Researchers should benefit from exploring insights into its implementation in the real world. The findings are useful not only to researchers and practitioners in the Vietnam Construction Industry(VCI) but also to participants in other developing countries in South East Asia. Since Korea has emerged as the first largest foreign investor in Vietnam, the results of this study may be also useful to participants in Korea.

Key Words: artificial neural networks(ANN), Neural networks, cost model, prediction, total construction cost, Vietnam.

1. Introduction

Estimating plays a key role in the successful completion of the construction projects. The correctness of estimates depends on the estimator's experiences. Experienced estimators rely on their personal expertise to incorporate the effect qualitative factors in their estimates (Ezeldin and Sharara 2006). Neural networks are the tools that help less experienced estimators to integrate such effects. Like human brain, neural networks learn from experiences, generalize from previous examples to new ones and abstract essential characteristics from inputs containing irrelevant data (Boussabaine 1996).

* 일반회원, Senior lecturer Division of Construction Engineering & Management Faculty of Civil Engineering HoChiMinh City University of Technology, 공학박사. luutruongvan@yahoo.co.kr

** 종신회원, 부경대학교 건설공학부 교수, 공학박사(교신저자) kims@pknu.ac.kr

A common feature found is that there is an exclusive concentration on the use of artificial neural networks to predict construction productivity of various operations. It seems that little effort has been spent to apply neural networks to apartment in developing countries, where the construction environment is risk due to the impact of the changes on direct costs. As a result, a distinct need has emerged for effective tools to help estimators working in developing countries better estimating their costs. This study is focusing on Vietnam, a developing country in the South East Asia and a new member of WTO. In Vietnam, estimators usually have encountered many problems in their works due to lack of cost information. Therefore the accuracy of cost estimates play a key role in cost management for owners. An effort is made to illustrate the use of neural networks for forecasting TCC of apartment projects in this local market.

The major objective of this study is to develop a NN-based model for estimating TCC of apartment projects in conceptual design phase. The results are useful not only to researchers and practitioners in the Vietnam Construction Industry but also to participants in other developing countries in South East Asia.

Korea has emerged as the first largest foreign investor in Vietnam, pouring over 4.2billion USD in 403 projects in 2007(VNA 2007). Therefore, the results are useful not only to practitioners and researchers in Vietnam but also to participants in Korea and

2. Previous works

2.1 Previous works

The neural network has been applied successfully in a variety of construction majors, including the estimate construction productivity for concrete form work tasks (Portas and AbouRiz 1997), the quantitative evaluation of the impact of multiple factors on concrete productivity (Sonmez and Rowings 1998), cost flow forecasting (Boussabaine and Kaka, 1998), the construction mark up estimation (Li et al. 1999), the prediction of earthmoving production (Shi 1999), the forecast of residential construction demand (Goh 1998 and 2000), the prediction of the level of organizational effectiveness in a construction firm (Sinha and McKim 2000), the improvement of construction qualifications in terms of the objectiveness (Lam et al. 2001), the processing the subjective information needed for crane type selection in a consistent manner (Sawhney and Mund 2002), the prediction of total construction cost (Emsley et al. 2002), the estimation of the acceptability of new form work systems (Elazouni et al. 2005). Recently, Ling and Liu (2004) have demonstrated the applicability of ANNs to predict the performance of design-build projects. Wilmot and Mei (2005) suggested the NN models of highway construction costs. Ezeldin and Sharara (2006) developed neural networks capable of predicting the productivity rates of forms assembly, steel fixing and concrete pouring. Ok and Sinha (2006) developed a non-linear analysis using neural network model for estimating construction productivity of dozer operations. Moreover, the use of neural network technique in cost prediction under the various impacts of subjective factors has appeared in several scientific journals and technical reports (Margaret et al. 2002; Mohamad et al. 2006).

2.2 Vietnam construction market

The purpose of this section is to highlight characteristics of Vietnam construction market(VCM) being one of features of the paper. (Luu et al. 2008) provided a brief description of salient characteristics of VCM. Those characteristics are:

(1) VCM heavily depends on foreign and government orders in a segment of industrial plants (i.e. processing plants, electricity power plants, petroleum and gas disposal facilities, and so on) and infrastructure projects, whereas private orders in other segments. Lacks of capital may be a barrier to domestic order from a private sector.

(2) The market is mainly controlled by large and medium construction firms. The proportions of the

firm in terms of size are: small (74%), medium (20%) and large (6%) (Luu et al. 2008). However, outputs of medium and large construction firms (i.e. net revenue) contributed 89% to total revenue of the construction sector (Luu et al. 2008).

(3) The construction order may differ from region to region. The construction orders are allocated around large cities such as Ha Noi, Ho Chi Minh city, and industrial zones because of the differences in economic growth.

(4) The central government and local authorities are major customers for large projects. As a result, close and cooperative relation with local authorities is essential to achieve more orders.

(5) The construction demand is sharply increasing due to the promotion of industrialization and the official development assistant (ODA). The competitiveness in VCM is greatly influenced by the capability of capital supply. Therefore, it strongly requires construction firms with financial capacity.

Moreover, the salient points of the apartment sector are: (1) high concentration on high-class apartment projects; (2) imperfect market of the apartment segmentation; (3) very few efforts to invest in the segmentation of house for rent; (4) very high demand of low-cost apartment projects; and (5) intensive involvement of local authorities in the early phases of apartment projects.

The Vietnam construction industry intensively needs funds to meet rapid growth in apartment projects (Kim et al. 2008). However, the financial capacity of Vietnamese investors cannot always meet the capital demand. As a result, Vietnamese developers prefer investments in high-class apartment projects to others. Moreover, preferential treatment for state construction firms and imperfect information on urban planning also caused imperfect market of the apartment segmentation. Since VCM is probably high risks, there are very few efforts to invest in the segmentation of house for rent because a low profit and long pay-back period. Vietnam is still a relatively poor country with USD 726 per capita at the market exchange rate (2006 estimate) (Kim et al. 2008). As a result, the majority of people seem to prefer low-cost apartments to high-class apartments because of limitation of ability to pay. In Vietnam, the central government reserves for itself land ownership whereas people have only the right of land use. Therefore, intensive involvement of local authorities in the early phases of apartment projects is common. This may cause corruption and bureaucracy in issuing land use certificates.

In addition, VCM has faced many problems such as complexity of legal and institutional framework, lack of capable consultants and domestic contractors for handling large projects(Long et al. 2004), poor change management, bureaucracy of several local authorities, fraudulent practices, and kickbacks (Luuet al. 2008).

3. Developing the Neural Network model

3.1 Conceptual framework

This study focuses on a neural network based approach for forecasting the preliminary construction cost of apartment projects. A systematic methodology was implemented in order to achieve objectives of the study (Figure 1). It comprises of main four steps: (1) identifying input variables and choose a proper neural architecture; (2) collecting data for training and testing; (3) training and testing the neural network model; (4) applying the NN-based model for predicting the TCC.

3.2 Identifying input variables

Literature review was a first step to develop a survey form that helps to identify factors that affecting TCC of apartment projects. In order to fit into conditions in the VCI, the preliminary questionnaire was tested. Fourteen experts, who had at least 10 years of experience of building projects, were involved in the pilot test to revise the questionnaire. The final questionnaire was distributed to construction professionals involved in apartment projects in Ho Chi Minh city. A total of 91 questionnaires were collected to identify significant factors which can be

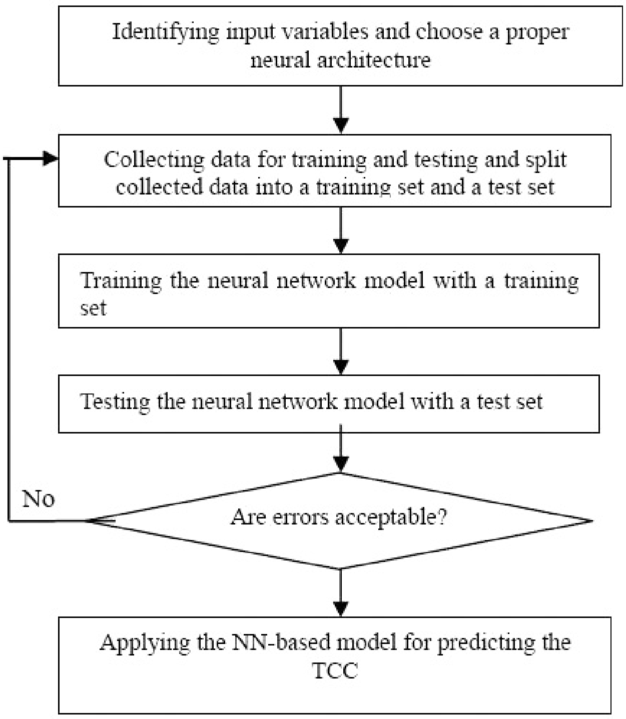

Figure 1: The conceptual framework for developing the NN model

Based on characteristics of a set of available data, seven input variables were identified as input variables categorized as scope variables and material price related variables (Table 1)

Table 1: Classification of input variables

Scope variables	Material variables price related
Prank of the project	
Gross Floor Area(GFA)	
Number of floors Year	
of construction	
	Petrol price(average price)
	Steel price(average price)

GFA, number of storys, petrol price, steel price, and cement price are real figures. In addition, those prices are deflated for adjusting the data to the same basis. The remaining variable is presented by a single input. According to Vietnamese Construction Law, construction projects are classified by four distinct ranks which is an order from "1" up to "4". This rank is in as much as an increase in preliminary construction cost. A single input is appropriate for "Rank of the project" variable because its value increases then it must represent an increase in some factors that affects the outcome of the model. Moreover, in order to be suitable for neural network (NN) processing, the study adopted the formula for scaling suggested by Hegazy and Ayed (1998). Prices of steel, cement and petrol is scaled to a range from -1 to 1.

There are numerous network architectures, therefore the following principles are considered for selecting the appropriate neural network.

· The number of input nodes are determined by the number of independent variables used in this model;
· The number of hidden nodes can be based on 75% of the number of input nodes;
· The number of hidden layers should be minimized to avoid longer training period for the neural network (Ezeldin and Sharara 2006);
· The number of neurons should be sufficient for the network to converge yet they should not be exaggerated to make the network memorize (Ezeldin and Sharara 2006).

Based on above principles, a three-layered feed-forward network was chosen. Moreover, two other networks, i.e. radial basis functions and generalized regression neural networks, were experimented to assess the best method. Based on R^2 and mean absolute percentage error (MAPE), three-layered feed-forward network resulted in the best performance. The formulation of MAPE can be found in the section 4.

This model consists of an input layer with 7 nodes, and an output layer with one node. Since one hidden-layer NN is suitable for most application in construction (Hegazy et al. 1994), the model fixed on one hidden-layer NN.

As a rule of thumb, "the number of hidden nodes was set as on-half of the total input and output nodes, as heuristically suggested in the literature (Hegazy et al. 1994)". The study fixed on five hidden nodes. However, "designing the network architecture and setting its parameters required some trial and error (Hegazy et al. 1994)”, the network was experimented with three, five, seven and ten hidden nodes. Of theses alternatives, the experiment with five hidden nodes offered the best performance, in terms of the associated value of MAPE.

3.3 Collecting data

The data use for training was actual data set of 14 apartment projects. Only data for projects that were absolutely completed were used for training the neural network. All projects used in this study are considered as a medium size. In Vietnam, there is no official definition of a medium-sized construction project because the Construction Law in Vietnam classifies a construction project as type 1, 2, 3 and 4. This Law did not classify a construction project as small, medium and large. Therefore, the research defines a medium construction project as a project with a total construction cost from 3 billion VND (USD1.00?VND16922) to 170 billion VND. The duration of completed projects ranged from 12 months to 18 months. The data were collected from Department of Planning and Investment of Ho Chi Minh city in the past seven years. Other data, such as steel, cement and petrol price were collected from Statistical Office in Ho Chi Minh city. Figure 2 illustrates the prices of steel, cement and petrol during 2000~2006. A data set of other five projects was used in the testing and verification of the model.

Since a large proportion of apartment projects in Vietnam are located on Ho Chi Minh city (Kim et al. 2008), the study selected the sample including apartment projects located on that city. Moreover, Department of Planning and Investment of Ho Chi Minh city deems that detailed tender documents are confidential. Therefore, some contractors were contacted to obtain specific information on apartment projects used in the model. Detailed information on the sample can be found in Appendix 1.

3.4 Training the NN

Since back-propagation training is one of the most common methods for training neural networks given historical data (Hegazy and Ayed 1998), this method was used for training the model. A commercial software “MATLAB” was used as a tool for the NN training. MATLAB was chosen because of its easy of use and speed training. In addition, in order to evaluate the best algorithm to training, a number of different algorithms were tried: quasi-Newton, gradient descent, Levenberg-Marquardt, and gradient descent with momentum. This resulted in Levenberg-Marquardt algorithm as the most effect. The parameters used to train the neural network is shown in Figure 3.

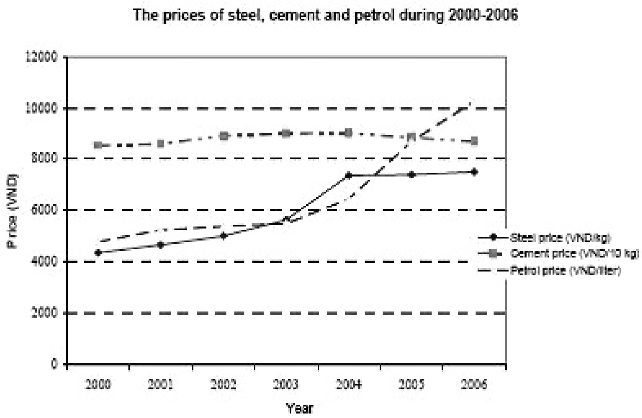

Figure 2: The prices of steel, cement and petrol during 2000-2006

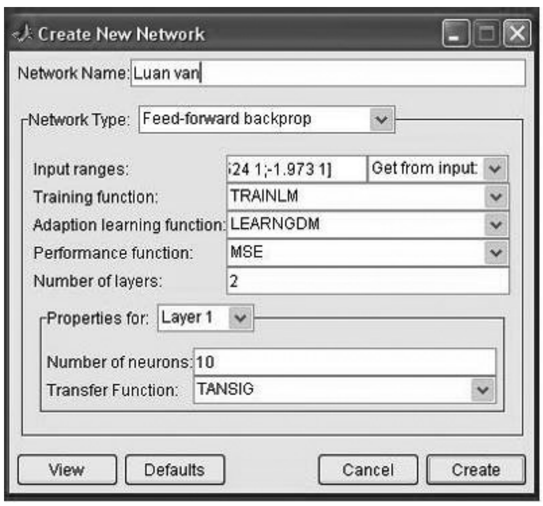

Figure 3: MATLAB parameters for training the model

In order to evaluate the best transfer function, three different function types were experimented: linear, logistic, and tanh. The tanh function offered the best results. Hegazy and Ayed (1998) reported that the tanh function is appropriate for their model.

The first 14 cases were used in training and the last five cases were used for testing. In general, back-propagation training adapts a gradient-descent approach of adjusting the neural network weights. The data of thousands of training cycles (called epochs in MATLAB) present during a neural network is trained. After each cycle, the error between the NN outputs and the actual outputs are propagated backward to adjust the weights in a manner that is mathematically guaranteed to converge (Hegazy and Ayed 1998). The process of training considered 5,000 epochs in testing because the appropriate training and acceptable converge were achieved. The mean square error (MSE) for a second train is 4.85403x10-16(Figure 4). This value is acceptable for generalization from data.

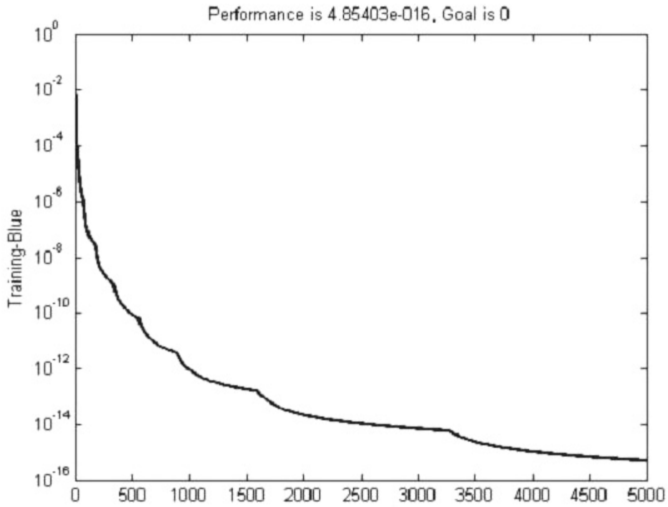

Figure 4: MSE versus the epochs

4. Results and discussion

Based on the weight matrix resulted from MATLAB, a computerized tool, namely TCC Prediction Tool(TPT), has been developed in order to predict the TCC for apartment projects. The tool is user-friendly and designed to facilitate the manipulation via Visual C++

software. Figure 5 illustrates a part of this tool.

The best way to evaluate the forecast model is comparison between its actual and predicted results. The tool was applied to 5 real-case apartment projects. All projects are apartment projects, completed and located in Ho Chi Minh City, a largest city in Vietnam. Table 2 shows prediction results of those real cases. The above pilot test of the tool has proved the validity of the proposed model in this study.

The study adopted three relative measures of accuracy dealing with percentage errors pointed out in Goh(2000) to compare the forecasting results of the NN-based model. Those measures are:

· Percentage error (PE) :

$$PE_t = \frac{X_t - F_t}{X_t} \times 100\% \qquad (1)$$

where Xt is the actual value of cast t : and Ft is the forecast value of case t

· Mean percentage error(MPE)

$$MPE = \sum_{i=1}^{n} \frac{PE_i}{n} \qquad (2)$$

where n is the number of forecasts

Table 2: Prediction results of five real-case apartment projects

Project	Prediction results of construction costs by the tool (VND)	Actual construction costs (VND)	Prediction results of construction costs by other methods (VND)*	Percentage error 1 (PE1) ((3)-(2)) / (3)	Percentage error 2 (PE2) ((3)-(4)) / (3)
1	2	3	4	5	6
A	30,182,900,000	28,510,000,000	33,562,700,000	-5.90%	-17.70%
B	12,952,600,000	14,062,000,000	15,768,300,000	7.90%	-12.10%
C	6,183,310,000	6,503,200,000	7,854,100,000	4.90%	-20.80%
D	17,983,400,000	15,104,800,000	16,189,500,000	-19.10%	-7.20%
E	14,820,300,000	15,527,600,000	13,444,600,000	4.60%	13.40%
MPE				-1.50%	-8.90%
MAPE				8.50%	14.20%

· Mean absolute percentage error(MAPE)

$$MAPE = \sum_{i=1}^{n} \frac{|PE_i|}{n} \qquad (3)$$

where $|PE_i|$ is the absolute value of the percentage error of case I.

As shown in Table 2, percentage error 1 (PE1) describes the variances between actual TCC of projects and predicted results from the ANN-based model, while percentage error 2 (PE2) represents the variances between actual TCC and the former TCC from other methods. The forecasting accuracy of the model may be reflected in its MAPE value (Goh 2000). Since the MAPE value of the ANN-based model consistently fall within the acceptable limit of 10% (Goh 2000), the capability of ANN in predicting TCC of apartment projects has been confirmed and reinforced in this study.

In order to access the model value, two other methods were tried: multiple linear regression (MLR) and genetic algorithm(GA). Of these alternatives, i.e. ANN, MLR and GA, the ANN method provided the best performance, in terms of the value of MAPE. Therefore, it can be concluded that the ANN approach offered the best model for predicting TCC of apartment projects corresponding to sample investigated. However, the ANN-based model in this study may apply to apartment projects, which have the characteristics to be similar with apartment projects in the sample.

5. Conclusions

The development of models to estimate the construction cost of apartment projects has gained considerable interests from researchers as well as practitioners. The best prediction of construction costs will provide assistance to project managers in taking proactive actions to secure project success. An ANN model can be used to assist the construction professionals in predicting, updating project cost (Boussabaine 1996). The main purpose of this paper is to develop an ANN-based model

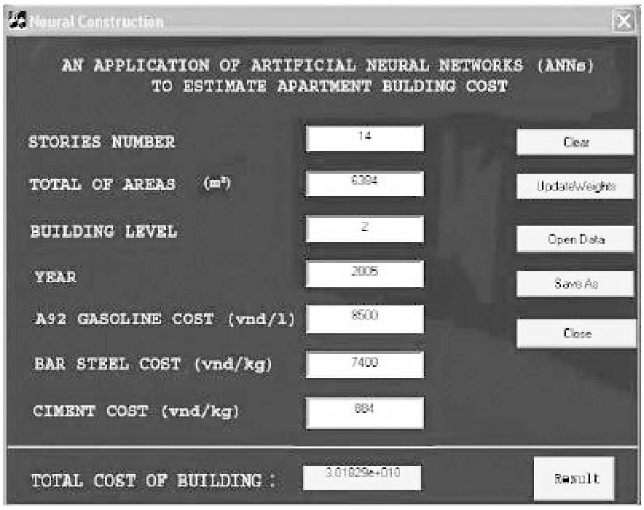

Figure 5: TCC Prediction Tool (TPT): Prediction results screen shot

to predict construction costs for apartment projects in Vietnam. The comparison of predicted and actual result suggested that the ANN-based model has potential to improve the cost estimation model for apartment projects. Although the proposed model is not validated in a rigorous way, the ANN-based model is useful for both practitioners and researchers. It facilitates systematic predictions in early phases of construction projects.

Practitioners are more proactive in estimating construction costs and making consistent decisions in initial phases of apartment projects. Researchers should benefit from exploring insights into its implementation in the real world. The results of the tool are conclusive so that other project stakeholders can analyze an apartment project in an intensive way. Moreover, the results confirmed the strength of ANN in predicting TCC of apartment projects.

The outcomes of our research are models for predicting the total construction cost of different building projects. Therefore, the government authorities can use the outcomes of our research to obtain the basis for the price evaluation. Moreover, concessionaires can apply the prediction models as a criterion to benchmark their estimated construction costs. Eventually, the outcomes of this research provide better information for the government authorities as well as concessionaires to negotiate, and mitigate a high possibility for disputes, which is an inherent characteristic of building projects

However, the limitations of the study are listed hereinafter

- The prediction tool should be improved in terms of continuous updating the weight matrix by adding more data of completed apartment projects
- The weight matrix embedded in the tool should be validated in a more rigorous way
- Other input variables are needed to improve the prediction tool.

Finally, further research seeks to develop a proper and realistic model for accurately estimating construction costs of apartment projects

References

1. Boussabaine, A.H. (1996). "The use of artificial neural networks in construction management: a review". Construction Management and Economics, 14(5), pp. 427-436.
2. Boussabaine, A.H., and Kaka, A.P. (1998) A neural networks approach for cost flow forecasting, Construction Management and Economics, 16(4), pp. 471-479.
3. Elazouni, A.M., et al. (2005). Estimating the acceptable of new formwork systems using neural networks. Journal of Construction Engineering and Management, ASCE, 131(1), pp. 33-41
4. Emsley, M.W., et al. (2002) Data modelling and the application of a neural network approach to the prediction of total construction cost. Construction Management and Economics, 20(6), pp. 465-472
5. Ezeldin, A.S., and Sharara, L.M. (2006). "Neural networks for estimating the productivity of concrete activities". Journal of Construction Engineering and Management, ASCE, 132(6), pp. 650-656.
6. Goh, B.H (1998) Forecasting residential construction demand in Singapore: a comparative study of the accuracy of time series, regression and artificial neural network techniques. Engineering, Construction and Architectural Management, 18, pp. 261-275.
7. Goh, B.H (2000) Evaluating the performance of combining neural networks and genetic algorithms to forecast construction demand: the case of Singapore residential sector. Construction Management and Economics, 18(2), pp. 209-217.
8. Hegazy, T., et al. (1994). "Developing practical neural network application using back propagation". Journal Microcomputers in Civil Engineering, 9(2), pp. 145-159.
9. Hegazy, T., and Ayed, A. (1998) Neural network model for parametric cost estimation of highway projects. Journal of Construction Engineering and Management, ASCE, 124(3), pp. 210-218.
10. Kim, Y.M., et al. (2008). Cause of construction delays of apartment construction projects: Comparative analysis between Vietnam and Korea. Korean Journal of Construction Engineering and Management, 2008(10), pp. 214-226.
11. Lam, K.C., et al. (2001) A fuzzy neural network approach for contractor prequalification. Construction Management and Economics, 19(2), pp. 175-188.
12. Li, H., et al. (1999) ANN-based mark-up estimation system with self-explanatory capacities. Journal of Construction Engineering and Management, ASCE, 125(3), pp. 185-189.
13. Ling, F.Y.Y., and Liu, M. (2004) Using neural network to predict performance of design-build projects in Singapore. Building and Environment, 39, pp. 1263-1274
14. Long, N.D., et al. (2004). Large construction projects in developing country: a case study from Vietnam. International Journal of Project Management, 22(7), pp. 553-561.
15. Luu, T.V., et al. (2008). Performance measurement of construction firms in developing countries. Construction Management and Economics, 26(4), pp. 373-386.
16. Margaret, W.E., et al. (2002) Data modelling and

the application of a neural network approach to the prediction of total construction costs. Construction Management and Economics, 20, pp. 465-472.

17. Mohamad Zin, R., et al. (2006) Predicting the performance of traditional general contract projects: a neural network based approach. Proceedings of 6th APSEC, 5-6 September 2006, Kuala Lumpur, Malaysia, pp. C-78-C-86.
18. Ok, S.C. and Sinha, S.K. (2006) Construction equipment productivity estimation using artificial network model. Construction Management and Economics, 24(10), pp 1029-1044.
19. Portas, J. and AbouRiz, S. (1997) Neural network model for estimating construction productivity. Journal of Construction Engineering and Management, ASCE, 123(4), pp. 399-410.
20. Sawhney, A., and Mund, A. (2002) Adaptive probabilistic neural network-based crane type selection system. Journal of Construction Engineering and Management, ASCE, 128(3), pp. 265-273.
21. Shi, J.J. (1999) A neural network based system for predicting earthmoving production. Construction Management and Economics, 17(4), pp. 463-471.
22. Sinha, S., and McKim, R.A. (2000) Artificial neural network for measuring organizational effectiveness. Journal of Computing in Civil Engineering, ASCE, 14(1), pp. 09-14.
23. Sonmez, R. and Rowings, J.E. (1998) Construction labor productivity modelling with neural networks. Journal of Construction Engineering and Management, ASCE, 124(6), pp. 498-504
24. VNA. (2007). Republic of Korea is still a first foreign investor in Vietnam, Vietnam News Agency, retrieved from website: http://www.vnagency.com.vn/TrangChu/VN/tabid/58/itemid/229893/Default.aspx
25. Wilmot, C.G., and Mei, B. (2005) Neural network modelling of highway construction costs. Journal of Construction Engineering and Management, ASCE, 131(7), pp. 765-771.

논문제출일: 2008.03.12
심사완료일: 2009.01.05

<Appendix 1> The data set of 14 projects for training

Apartment project	No. floors	GFA (m2)	Project rank	Project start	Project finish	Location	Total construction cost (VND)
1	12	9259	2	2000	2001	HCM city	24,274,596,203
2	12	21363	2	2002	2004	HCM city	54,377,505,193
3	25	41289	1	2002	2004	HCM city	165,423,049,019
4	12	18680	2	2003	2005	HCM city	65,379,591,922
5	5	3819	3	2003	2004	HCM city	7,614,996,967
6	12	11990	2	2003	2004	HCM city	39,579,000,000
7	12	10313	2	2003	2004	HCM city	42,781,413,000
8	5	1508	3	2003	2003	HCM city	3,007,261,000
9	11	16838	2	2004	2006	HCM city	37,569,769,727
10	14	6384	2	2005	2006	HCM city	28,510,000,000
11	9	5482	3	2005	2006	HCM city	17,951,921,743
12	5	2413	3	2005	2005	HCM city	8,447,875,363
13	15	13350	2	2005	2006	HCM city	46,724,764,835
14	15	11857	2	2006	2007	HCM city	45,057,739,755

Note: HCM: Ho Chi Minh city; GFA: Gross floor area; Project rank: rank of the project based on the Vietnamese Government's Decree No. 16/2005/ND-CP of February 07, 2005; 1 USD = 16,992 VND (13-December-2008).