AUGMENTED INTEGRATED FUZZY NEURAL NETWORK LEARNING MODEL IN STRUCTURAL ENGINEERING

Shih-Lin Hung and Jan, C.C.

Department of Civil Engineering, National Chiao Tung University
1001 Ta Hsueh Road, Hsinchu, Taiwan 300, R.O.C.

Abstract: Solving engineering problems is a creative, experiential process. An experienced engineer generally solves a new problem by recalling some similar instances examined before and applying what he learned from the present problem, through adaptation or synthesis. According to such a method, the IFN learning model was developed and implemented as a computational model for problem solving. This model has been applied to design problems involving a complicated steel structure. Computational results indicate that the IFN model can learn the complicated problems within a reasonable computational time owing to its simplicity. The learning performance of IFN, however, relies heavily on the values of some working parameters selected on a trial and error basis. In this work, we present an augmented IFN learning model by integrating a conventional IFN learning model with two novel approaches: a correlation analysis in statistics and a self-adjustment in mathematical optimization - to facilitate the search for appropriate working parameters in the conventional IFN. The problem of arbitrary trial and error selection of the working parameters is avoided in the augmented IFN learning model.

Keywords:Artificial Neural Networks, Supervised Learning Models, Unsupervised Fuzzy Learning Model, Structural Design.

1. TNTRODUCTION

Solving engineering problems, such as analysis and design problems, is a creative, experiential process in which the experiences and combined knowledge of engineers serve as resources. An experienced engineer generally solves a new problem in the following stages. First, he or she recalls some similar instances that have been solved before while properly considering the functional requirements of those instances. Then, the engineer attempts to derive the solution from these similar instances through adaptation or synthesis. After the problem is solved, the new instance is then stored in her/his memory as an additional knowledge resource for solving further problems in the future.

The above stages can be implemented as a computational model for problem solving that utilizes a case base of previously solved problems when solving a new problem. In symbolic artificial intelligence (AI), case-based reasoning [1] is an effective means of facilitating computer program development that attempts to solve problems by directly accessing the case base. Artificial neural networks (ANNs), on the other hand, constitute a different AI approach that has made rapid advances in recent years. Such networks have the ability to develop, from training instances, their own solutions to a class of problems [2-5].

Recently, authors [6][7] presented an integrated fuzzy neural network (IFN) learning model in structural engineering. The IFN learning model combined a novel unsupervised fuzzy neural network (UFN) reasoning model with a supervised neural network learning model using the adaptive L-BFGS learning algorithm [8]. The IFN learning model was applied to steel beam design problems. That work contended that the IFN learning model is a robust and effective ANN learning model. In addition, the IFN model can interpret a large number of instances for a complicated engineering problem within a reasonable computational time owing to its simplicity in computation. However, the performance of the IFN learning model is heavily affected by some working parameters that are problem-dependent and obtained via trial and error.

Herein, we present a more effective neural network learning model, called the augmented IFN learning model, by integrating a conventional IFN learning model with two newly developed approaches: a correlation analysis in statistics and self-adjustment in mathematical optimization. The first approach,

correlation analysis in statistics, is employed to assist users in determining the appropriate working parameter to be used in the fuzzy membership function. The second approach, self-adjustment in mathematical optimization, is used to obtain appropriate weights, systematically, for each decision variable required in the input of training instances. The augmented IFN learning model proposed herein is applied to the preliminary design of steel structure buildings.

2. REVIEW OF IFN LEARNING MODEL

The IFN learning model combines two sub-ANN learning models [6]. One is a novel unsupervised fuzzy neural network (UFN) reasoning model: a single-layered laterally-connected network with an unsupervised competing learning algorithm. The other is an offline assistant model: a supervised neural network learning model with the adaptive L-BFGS learning algorithm [8]. The IFN learning model is schematically depicted in Fig. 1.

The learning stage in the IFN model is performed in two sub-ANN models concurrently. First, the offline assistant supervised neural network model is trained, based on the adaptive L-BFGS supervised learning algorithm. In the UFN reasoning model, however, the learning process simply involves selecting appropriate working parameters for the fuzzy membership function and weights for each decision variable in the input.

After learning in the UFN and in the assistant supervised learning model is completed, any new instance X can be solved via the IFN learning model. The reasoning in UFN is performed through a single-layered laterally-connected network with an unsupervised competing algorithm and it is implemented in three steps. The first step involves searching for some instances from the instance base that resemble the new instance X according to their inputs.

$$d_{ij}=diff\left(U_{i,i},U_{j,i}\right)=\sum_{m=1}^{M}\alpha_{m}(w_{i}u_{i}^{m}-w_{j}u_{j}^{m})^{2} \qquad (1)$$

The second step entails representing the relationships among the new instance and its similar instances, with d_{Xj} less than $R_{\max}$.

$$\mu\left(d_{Xj}\right)=\begin{cases}0 & if \ d_{ij}\geq R_{\max}\\ \dfrac{R_{\max}R_{\min}-R_{\min}d_{ij}}{(R_{\max}-R_{\min})d_{ij}} & if \ R_{\min}<d_{ij}<R_{\max}\\ 1 & if \ d_{ij}\leq R_{\min}\end{cases} \qquad (2)$$

The final step involves generating the output X_o vector of instance X by synthesizing the outputs of its similar instances according to their associated fuzzy membership values using center of gravity (COG) or mean of maximum (MOM) methods [6][7].

The reasoning process of the UFN depends on determining the degree of similarity among X and U_j. Consequently, no solution can be generated by the

UFN reasoning model, if the new instance entirely differs from all instances in the instance base, e.g. all d_{Xi} are greater than $R_{\max}$. In addition, using inappropriate working parameters would allow for the possibility that no similar instances can be derived. For the above issues, the under-trained adaptive L-BFGS supervised neural network is used as an assistant system to generate an approximate output for the new instance.
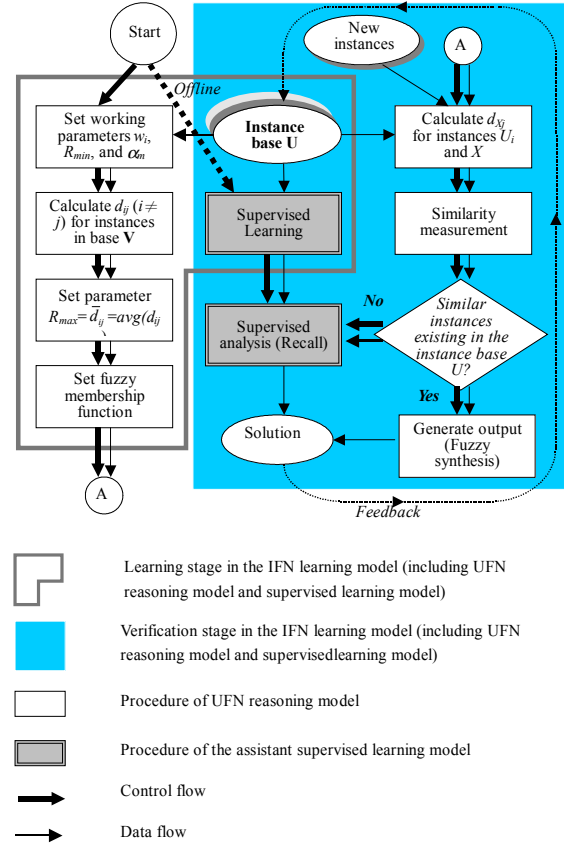

Figure 1   Conventional IFN learning model.

3. AUGMENTED IFN LEARNING MODEL

In the conventional IFN learning model, working parameters, such as w_j, α_m, $R_{\max}$, and $R_{\min}$, are selected subjectively by users and, generally, on a trial and error basis. Consequently, the learning performance is highly affected by these parameters, especially $R_{\max}$ and α_m. In this work, two novel approaches are employed for assisting the users to determine these parameters and weights systematically.

3.1 Correlation Analysis for $R_{\max}$ in Fuzzy Membership Function

In conventional IFU, the similarity measurement between two instances heavily depends on the value of parameter $R_{\max}$. A small value of $R_{\max}$ implies that a

strict similar relationship between instances is utilized. Consequently, most of the instances are sorted as dissimilar. Consequently, few similar instances to the new instance can be found and, ordinarily, no solution can be generated via the UFN reasoning model. On the other hand, a loose similar relationship is adopted under the case of a larger R_{max}. Accordingly, a large number of instances are taken to be "similar instances" and the solution generated via these similar instances, is inferior. Here, the linear correlation analysis in statistics is employed to facilitate the determination of appropriate value of R_{max} in the fuzzy membership function. The analysis is a process which aims to measure the strength of the association between two sets of variables that are assumed to be linearly related.

For the above instance base **U** with N instances, the correlation analysis in the fuzzy membership function is implemented in the following steps. The first step is to determine the degree of difference between any two instances in the base **U** using the aforementioned function of degree difference in Eq.(1). Hence, a total of ($C_2^N + N$) resembling samples, $S_{ij}(U_{i,o}, U_{j,o}, d_{ij})$, can be compiled. A resembling sample contains two instances' outputs ($U_{i,o}$ and $U_{j,o}$) and the corresponding degree of difference (d_{ij}). Thereafter, two arrays, A_t and B_t, can be assorted from resembling samples in the case of d_{ij} less than or equal to a prescribed value, say t. The elements in A_t and B_t are the first and second items, respectively, of these resembling samples. Next, the accumulative correlation coefficient, $Ac\_CORREL(A_t, B_t, t)$, is calculated for arrays A_t and B_t with the degree of difference less than or equal to t. Assume that for a total P resembling samples with d_{ij} less than a prescribed t, the arrays A_t and B_t can be denoted as

$$A_t = \{a_k \mid a_k = U_{i,o} \in S_{ij}, \ for \ d_{ij} \leq t\} = \{a_1, a_2, ..., a_p\}$$

$$B_t = \{b_k \mid b_k = U_{j,o} \in S_{ij}, \ for \ d_{ij} \leq t\} = \{b_1, b_2, ..., b_p\}$$

The value of the accumulative correlation coefficient equals

$$Ac\_CORREL(A_t, B_t, t) = \frac{Cov(A_t, B_t, t)}{\sigma_{A_t} \sigma_{B_t}} \quad (3)$$

$$Cov(A_t, B_t, t) = \frac{1}{p}\sum_{k=1}^{p}(a_k - \mu_{A_t})(b_k - \mu_{B_t}) \ \ s.t. \ d_{ij} \leq t \quad (4)$$

where σ_{A_t} and σ_{B_t} are standard errors of arrays A_t and B_t; μ_{A_t} and μ_{B_t} are the means of A_t and B_t. The formulas expressed in (3) and (4) represent the relationship between the accumulative correlation coefficient to any value of t. An accumulative correlation curve can be plotted as a function of t and $Ac\_CORREL(A_t, B_t, t)$. Note that the appropriate R_{max} equals a certain value of t such that instances in the instance base **U** have a certain degree of correlation.

Obviously, the smaller the t implies a larger accumulative correlation coefficient, indicating a strong relationship between the two arrays, e.g., the strongest correlation, $t=0$, between the two arrays refers to the case in which the instances in the two sets are identical and the value of $Ac\_CORREL(A_t, B_t, t)$ equals one. In such a case, no solution to a new instance can be generated via the UFN reasoning model except for when identical instances exist in the instance base. In order to avoid this issue here, we set $Ac\_CORREL(A_t, B_t, t)$ equal to 0.8 as the lower bound for similarity measurement. The value of t corresponding to this lower bound is adopted as the appropriate value of R_{max}.

3.2 Self-Adjustment for Selecting Weights α_m

Except for R_{max}, the selected weights α_m also significantly affect the learning performance for the conventional IFN. This occurrence has been investigated in the earlier work [7]. The learning results indicated that significant improvements were achieved as the weights were gradually updated via a basis of heuristic knowledge associated with learning problems. In this study, a more systematical approach - self-adjustment based on mathematical optimization - is adopted to facilitate the search for appropriate weights.

For the above instance base **U** with N instances, the self-adjustment approach can be briefly stated as consisting of the following steps. First, set up the corresponding working parameters, R_{max}, R_{min}, and w_j. where R_{max} is determined using the aforementioned correlation analysis approach; and where parameter R_{min} and weights w_j, are set as constants in this work. Meanwhile, weights α_m for each decision variable in the input are directly initialized as one. Then, based on these working parameters, the outputs for training instances are found via the UFN reasoning model. Then the error, E_i, between the computed and desired outputs, Y and $U_{i,o}$, for training instance U_i, is calculated. The system error, E, for a total N instances is then defined as half of the average sum of errors and denoted as

$$E = \frac{1}{2N}\sum_{i=1}^{N} E_i \quad (5)$$

$$E_i = \sum_{k=1}^{K}(y^k - o_i^k)^2 \quad (6)$$

where o_i^k and y^k are the kth items of data in desired and computed outputs, U_i and Y, respectively. Note that the system error is an implicit function of the weights α_m as $E(\alpha_m)$.

Weights α_m in the UFN reasoning model are adjusted to reduce the system error as much as possible. This goal can be achieved if a set of appropriate weights, α_m, are used. The problem, then, can then be considered as an unconstrained optimization problem; that is searching a set of

optimum weights by iteration to minimize the system error. In the mathematical optimization approaches, the conjugate gradient (CG) method has been proved an efficient means of solving the problem. The weights α_m are updated in each iteration, say the $(s+1)th$ iteration, as $\alpha_m^{(s+1)} = \alpha_m^{(s)} + \lambda \mathbf{d}^{(s)}$. The term $\square$ is step length and is set as a constant in this work. The search direction $\mathbf{d}^{(s)}$ is defined as

$$\begin{cases} \mathbf{d}^{(s)} = \mathbf{g}^{(s)} = -\nabla E(\alpha_m^{(s)}) & if \ s = 0 \\ \mathbf{d}^{(s)} = \mathbf{g}^{(s)} + \beta^{(s)} \mathbf{d}^{(s-1)} & if \ s > 0 \end{cases} \quad (7a,b)$$

where $\beta^{(s)} = \dfrac{(\mathbf{g}^{(s)})^T \mathbf{g}^{(s)}}{(\mathbf{g}^{(s-1)})^T \mathbf{g}^{(s-1)}}$, and $\mathbf{g}^{(s)} = -\nabla E(\alpha_m^{(s)})$

The iteration is terminated as the value of $\left\| \mathbf{g}^{(s+1)} \right\|$ or $\left| E(\alpha_m^{(s+1)}) - E(\alpha_m^{(s)}) \right|$ is sufficiently small. The term $\mathbf{g}^{(s)}$ is the negative gradient vector of function $E(\alpha_m)$. For simplicity, the superscript (s), denoted as the sth iteration, is ignored. Hereinafter, vector $\mathbf{g}$ can be derived as

$$\mathbf{g} = \frac{\partial E(\alpha_m)}{\partial \alpha_m} = \frac{1}{2N} \sum_{i=1}^{N} \frac{\partial E_i}{\partial \alpha_m} \quad (8)$$

$$\frac{\partial E_i}{\partial \alpha_m} = 2 \sum_k (y^k - o_i^k) \left(\sum_p f(d_{ip}) \right)^{-2} \times$$
$$\left(\sum_p o_p^k f'(d_{ip})(u_i^m - u_p^m)^2 \sum_p f(d_{ip}) \right.$$
$$\left. - \sum_p f(d_{ip}) o_p^k \sum_p f'(d_{ip})(u_i^m - u_p^m)^2 \right) \quad (9)$$

3.3 Augmented IFN Learning Model

Figure 2 schematically depicts the procedures of the augmented IFN learning model. Instead of using a constant working parameter R_{max} as in conventional IFN, the appropriate parameter is determined using correlation analysis. Meanwhile, the appropriate weights α_m are adapted during the self-adjustment process through a mathematical approach. The procedure of the augmented IFN learning model can be summarized as follows:

Self-organized learning phase:

Step 0: Train the adaptive supervised L-BFGS learning neural network model *offline*.

Step 1: Initialize parameter R_{max} as constant 10^{-5} and weight α_m for each decision variable on a heuristic basis or by trial and error.

Step 2: Calculate the degree of difference d_{ij} ($i{\neq}j$) among all instances in base $\mathbf{V}$.

Step 3: Determine the parameter R_{max} using correlation analysis and set $R_{max}=t$ such that $Ac\_CORREL(A_t,B_t,t)=0.8$.

Step 4: Set the fuzzy membership function, $\mu(d_{ij}) = f(R_{min}, R_{max}, d_{ij})$, defined in Eq.(2).

Step 5: Adjust the weight α_m for each decision variable in input using the self-adjustment approach.

Analysis phase (After the learning phase is completed):

Step 6: Present the new (unsolved) instance X to the UFN reasoning model and perform a similarity measurement between X and instance U_i in the base $\mathbf{U}$ using a single-layered lateral-connected competing network.

Step 7: If more than one similar instance is found in Step 6, generate the solution X_o for the new instance using fuzzy synthesis approaches in Eq. (4) or (5) and go to Step 9. Otherwise, go to next step.

Step 8: Compute the solutions via the under-trained assistant adaptive L-BFGS supervised learning model and go to next step.

Step 9: Feedback the new instance into the base $\mathbf{U}$. Meanwhile, further learning in the assistant supervised learning model is launched offline.
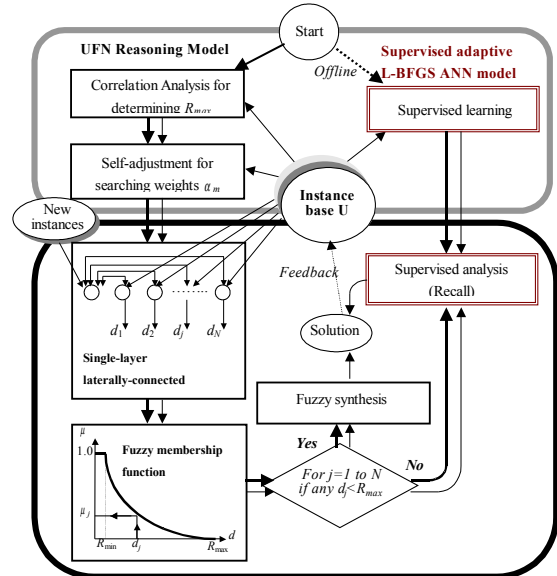

Figure 2　Augmented IFN learning model.

4. APPLICATIONS

In this work, the augmented IFN learning model is trained to learn how to implement the preliminary design of buildings satisfying the conditions of utility, safety, and economy in only one design cycle. In the complete design of a structure, the preliminary design stage is mainly a creative, experiential process that involves the choice of structure type, selection of the material, and determination of the sections of beams and columns in the structure. An experienced engineer is likely to carry out this stage more quickly than does an inexperienced one. The basic configuration of the structure at this stage should satisfy the specified, such as LRFD [9] for steel structures. To satisfy the prescribed constrains and achieve minimum

expenditure for materials and construction, this stage becomes a looped optimization decision-making process. Hence, a good initial development of the basic form, with sections of beams and columns satisfying the aforementioned constraints, will reduce the number of redesign cycles. The present example involves a complete design structure which satisfies the conditions that the service loads should not exceed the strength of the members; the drifts should be within the prescribed limits, and the structure should be economical in material (e.g. minimum weight), construction, and overall cost. For simplicity, only regular buildings with a rectangular layout, such as most factory buildings, are considered herein. Also, the beams in every floor have the same sectional dimensions, as do the columns.

416 instances are used in this example. They are randomly divided into 380 training instances and 36 verification instances. Seven decision variables are used as inputs to determine the sections of beams and columns of a building that satisfies the given specifications. The seven decision variables and their limits are described as following:

1. Number of stories = [9, 10, 11,12,13,14,15].
2. Bay length in long span direction (*X* direction)= 9 to 12 meters.
3. Bay length in short span direction (*Y* direction)= 6 to 9 meters.
4. Number of bays in both directions = [3, 4, 5].
5. Seismic zone coefficient = [0.18, 0.23, 0.28, 0.33].
6. Live load (kgw/m^2) = 200 to 350.
7. Wall load (kgw/m) = [100, 200].

Other corresponding decision variables used in the stage of preliminary design are assumed to be constant. Here, a building with three groups of steel elements in both beams and columns is considered. The three groups are upper, medium, and lower. An instance contains seven decision variable inputs and six data items as outputs.

The parameters R_{min} and w_i are set as constant 10^{-5} and one, respectively. The weights α_m, however, are initialized as [1, 1, 1, 1, 1, 1, 1] for the seven decision variables. Using these parameters and weights, correlation analysis in the augmented IFN learning model is performed first to determine the working parameter R_{max}. With the value of $Ac\_CORREL(A_t,B_t,t)$ equal to 0.8, the values of $t(= R_{max})$ are obtained. They are 0.12 for beams and 0.178 for columns, respectively.

After the fuzzy membership function is defined, the self-adjustment approach is then launched to obtain the adequate weights α_m for each decision variable in the input. The weights α_m for each decision variable are updated from [1, 1, 1, 1, 1, 1, 1] to [1.471, 1.369, 0.416, 0.008, 1.104, 0.825, 0.513] for beams and to [1.106, 0.997, 0.542, 0.009, 1.232, 0.999, 0.997] for columns. Interestingly, the weights of the fourth decision variable in beams and columns are both self-adjusted close to zero. This observation indicates that this decision variable (number of bays in both directions) is insignificant in the input. Consequently, this decision variable can be neglected. The 36 verification instances are used to verify the learning performance of the augmented and conventional IFN learning models, respectively. Notably, the augmented IFN is verified on the basis of the newly adjusted weights α_m with the adequate R_{max}. The working parameters and weights used in the conventional IFN, however, are selected on a trial and error basis. Table 1 summarizes the learning results for thirty-six verification instances. According to this table, the average percentage errors for beams and columns are 13.81 and 9.36 for the conventional IFN learning model. However, these errors are reduced to 6.17 and 6.1 for beams and columns, respectively, for the augmented IFN learning model. The augmented IFN learning model significantly improves in terms of learning.

TABLE 1. Results of Preliminary Structural Design Problem

Error (%)	Number of items in output for verification instances			
	Conventional IFN Learning Model		Augmented IFN Learning Model	
	Beam	Column	Beam	Column
0~5	12	43	52	51
5~10	29	29	39	31
10~15	21	22	13	20
15~20	29	10	3	6
> 20	17	4	1	0
Average Error	13.81 %	9.36 %	6.10 %	6.17 %

Note: Each instance has six items, three beams and columns, in output.

5. CONCLUUSIONS

This work presents an augmented IFN learning model by integrating two newly developed approaches into a conventional IFN learning model. These approached are a correlation analysis in statistics and self-adjustment in mathematical optimization, which collaboratively enhance the learning capability of the conventional IFN. The augmented IFN learning model proposed herein is applied to engineering design problem of preliminary design of stell structures. Based on the results of this work, we can conclude the following:

1. The problem of arbitrary trial and error selection of the working parameter (R_{max}), in fuzzy

membership function, encountered in the conventional IFN learning model is avoided in the newly developed augmented IFN learning model. Thus, the new learning model provides a more solid systemic foundation for IFN learning than the conventional IFN learning model.

2. In the conventional IFN learning model, the weights α_m, denoting the importance of the mth decision variable in the input, are set on a trial and error basis. This problem is avoided by the newly developed learning model. Instead of an assumed constant, the appropriate weights are determined through the self-adjustment approach in mathematical optimization. Therefore, the augmented IFN learning model provides a more solid mathematical foundation for neural network learning.

3. For each training instance, decision variables in the input are generally selected subjectively by users. As a result, some trivial decision variables may be adopted in the input for some complicated examples. Based on the self-adjustment approach, the importance of a decision variable in an input can be derived systematically. Therefore, insignificant or redundant decision variables in the input can be neglected.

Acknowledgments

REFERENCES

[1] Carbonell, J. G. (1981). "A computional model of analogical problem solving." *Proceedings of the Seventh International Joint Conference on Artificial Intelligence* 1: 147-152.

[2] Adeli, H., and Hung, S. L. (1995). *Machine Learning -- Neural Networks, Genetic Algorithms, and Fuzzy Systems*. John Wiley & Sons, New York, N.Y.

[3] Ghaboussi, J., Garrett, J. H., and Wu, X. (1991). "Knowledge-based modeling of material behavior and neural networks." *J. Engrg. Mech.*, 117(1), 132-153.

[4] Gunaratnam, D. J., and Gero, J. S. (1994). "Effect of representation on the performance of neural networks in structural engineering applications." *Microcomputers in Civ. Engrg.*, 9(2), 97-108.

[5] Hajela, P. and Berke, L. (1991). "Neurobiological computational modes in structural analysis and design." *Computers and Struct.*, 41(4), 657-667.

[6] Hung, S.L. and Jan, J.C. (1997), "Machine learning in engineering design- An unsupervised fuzzy neural network learning model," *Proceedings of Intelligent Information Systems (IIS'97), Grand Bahama, December 8-10, IEEE Computer Society, California, pp.156-160.*

[7] Hung, S.L. and Jan, J.C. (1999), "Machine learning in engineering analysis and design- An integrated fuzzy neural network learning model," *Computer-Aided Civ and Infrastructure Engrg.,* 14(3), 207-219.

[8] Hung, S. L., and Lin, Y. L. (1994). "Application of an L-BFGS neural network learning algorithm in engineering analysis and design." *Proceedings of The Second National Conference on Structural Engineering- Vol. III*, Nov. 4-6, Nantou, Taiwan, R.O.C., 221-230 (in Chinese).

[9] *Manual of steel construction – load & resistance factor design* (1994). America Institute of Steel Construction, Inc., Chicago, IL.