2003 Pacific Conference on Earthquake Engineering

Artificial neural network (ANN) modeling for earthquake damage detection in water distribution system

S. Takada

Department of Civil Engineering, University of Kobe, Kobe, Japan

N. Hassani & R. Rasti

Power and Water Institute of Technology (PWIT), Tehran, Iran

ABSTRACT: The artificial neural networks (ANN) as a non-parametric system identification approach present a robust and efficient way to simulate the nonlinear behavior of engineering systems. In this paper an artificial neural network, a general back error propagating perceptron, is used to detect damage in pipelines in water distribution systems which are involved in earthquakes. Since the pipes are buried underground, it is possible that their damage may not be found, even through extensive excavation. The failure point or points were obtained by using the amount of water discharge from the tanks. The states of probable failures then have to be computed. The amount of output discharge from the tanks is obtained by direct analysis for different states of pipe failure. Through a parametric study, different geometries, shapes, diameters and pressures of the water network are surveyed and the best network architecture for each case is obtained. The peak responses and phase delays are assumed to be the network outputs. Another feature is that the network can be operated in a supervised manner. The study shows the efficiency and capability of the ANN to model the observed nonlinear behavior.

1 INTRODUCTION

In a destructive earthquake considerable damage occurs in buried structures such as urban water supply systems. The preliminary action after an earthquake happens, is the repair of networks to allow the supply of water to the damaged area. The most difficult part of this work is to find the damaged locations.

In direct analysis of water networks, the pipe sizes, locations and elevations of reservoirs are determined in such a way that the discharges from the nodes or along the pipe segments and the necessary pressures in some nodes are provided.

In back analysis, water flow characteristics such as pressure, discharge, velocity and so on, are used as input data to detect the failed pipes and damaged locations. Since there are various types of failure and large amounts of input data, ANN has been applied. The identification of nonlinear problems has been receiving a growing amount of interest from researchers from diverse ranges of engineering. The conventional methods of identification, by either analytical or experimental methods, have some drawbacks.

Designing of a feasible method within time and money constraints, most of the time results in a limited number of parameters and thus a limited domain of application for these methods. These tools offer exciting attributes as the ability to adapt with highly nonlinear problems, fast parallel architecture and the capability of working as efficient sub-systems in large nonlinear super-systems. The most widely used type of neural networks in the field of engineering are the general feed-forward back error propagating perceptrons[1]. In this paper this kind of network is employed to simulate and model earthquake damage detection in water distribution system. The training data needed for constructing the network are generated from direct analysis of water networks.

2 THE NEURAL NETWORK MODEL

A neural network is an information-processing unit, originally intended to simulate the performance and characteristics of the human brain. The neural network is a computational technique able to learn characteristics of the introduced data and develop a generalization property. From a mathematical point of view, ANN is vector transformation, by which a vector is projected from n to m dimensions. Artificial neural networks are often named as black boxes, which is due to the fact that their functions are rarely evaluated. The ANN can be useful for complicated engineering problems. The ANN consists primarily of three basic elements: neurons, architecture of the net and a learning rule (Dayhoff 1990).

2.1 Back propagation neural network

There are different types of ANN. Back propagation is one of the most widely used of the neural network paradigms and has been applied successfully in applications studies in a broad range of areas. The back error propagating neural network is employed as the nonlinear identification approach in this study. These networks are introduced with three basic elements, which are nodes or units, layers or architecture and the activation function. The general configuration of these networks is shown in Fig. 1. The bottom layer of units is the input layer, the only units in the network that receive external input. The layer above is the hidden layer, in which the processing units are interconnected to layers above and below the top layer is the output layer.

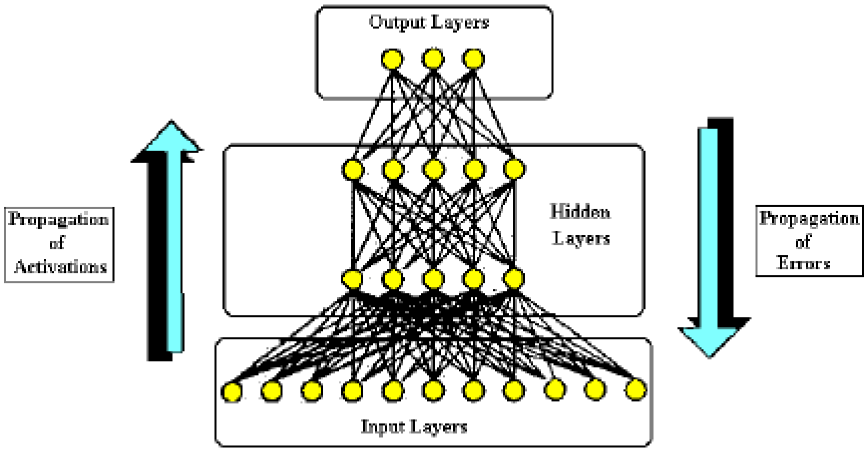

Fig. 1 Back propagation network topology (Dayhoff 1990).

2.2 Back propagation algorithm

The procedure of back propagation neural networks is briefly presented here:

- Set all weights, biases, weight modifiers and bias modifiers to random values in the desired ranges.
- Scale and present input vector to the input layer.
- Calculate input vector of the hidden layers:

$$Y_u^{in} = n_u + \sum_{h=1}^{H} x_h . w_{hu} \tag{1}$$

where Y = The input or output of the H hidden and K output layers; ? = Biases; $X = (x_1, x_2, ..., x_I)$ is the input vector with I members; and W are weights.

- Determine output vector using the transfer function:

$$Y_u^{out} = f(Y_u^{in}) \tag{2}$$

- Continue the procedure for all layers to obtain output vector.
- Calculate the error vector and total error to check for convergence:

$$d_u = (t_u - y_u).f'(Y_u^{in}) \tag{3}$$

where d is error in the layers; $T = (t_1, t_2, ..., t_U)$ is the target vector with U members.

$$E = \frac{1}{2}\sum_{u=1}^{U}(t_u - y_u)^2 \tag{4}$$

- Calculate weight and bias modifiers:

$$\Delta w_{hu} = a.d_u.y_h \tag{5}$$

$$\Delta n_{hu} = a.d_u \tag{6}$$

where a is learning rate.

- Modify weights and biases in the output layer:

$$w_{hu}^{new} = w_{hu}^{old} + \Delta w_{hu} \tag{7}$$

$$n_{hu}^{new} = n_{hu}^{old} + \Delta n_{hu} \tag{8}$$

2.3 Network training methodology

For suitable and successful training, the weighted relation of networks should not be close to unity. An appropriate solution, this is to use a deference between goal activity and computed activity for weighting adjustment and error reduction. This view is called “rule of Delta”. In the Delta Rule the weight varies between i and j provided the output is not zero or variation between output and real values. There are two methods for delta rule: the least mean square (LMS) and the threshold method.
The least mean square is applied when output units are non-linear. The total error is the sum of squares(Masri, Smyth, Chassiakos, Caughey & Hunter 2000).

$$E = \sum_p Ep = \sum_p \sum_i (tpi - upi)^2 \tag{9}$$

where E = error; p = training pattern; t_{pi} = target value; u_{pi} = real output.
In a back propagation neural network the error resulting from training and weights correction is called a loop, each loop contains 3 steps as listed:

1. Data obtained by the Network from input, and is able to determine the output units weight for training.
2. Output errors are computed and propagated back into the network.
3. Input modifyed and related to weights by analysis errors.

While a number of loops make up an epoch, network training may consists of hundred or even thousands of epoch. The Delta Rule minimizes the sum of LMS by gradient reduction. The results obtained by this method depend on the error values. If the error values are too small, the network converges and a result is obtained. Conversely, if the error values are large, the network diverges and no result can be obtained.
Five steps are to be considered in establishing a back propagation network:

1. The project to be dealt with ANN has to fit its activity.
2. Selection of architecture and topology (number of hidden layers, processing units and unit interaction)
3. Selection of the cases by which the ANN should be trained.
4. Real training of ANN in such a way those desirable results are generated in the output.
5. Behavior and strength of ANN for various cases.

Therefore, at first ANN architecture is selected then training patterns is determined and it is trained by Delta Rule. Finally, the network is tested which shows how the network responds to new input data.
The sampling data generated by the direct analyses of water distribution system in earthquake were stocked into a file and for each exemplar epoch a certain number of them were selected randomly

without repetition of any learning vector. Batching of concurrent inputs is computationally more efficient than sequential inputs and this is a reason for using epochs and total computation of errors.

In order to speed up the training phase some key ideas were considered. First was the online weight updating which was performed after the presentation of each exemplar. This will act as a random source of perturbation to the gradient descent and help the network to get away from being stuck in a local minimum and converge to the absolute minima.

Generally, in order to simulate a nonlinear relationship between inputs and outputs, a nonlinear activation or transfer function in needed. These functions show that how an ANN defines a system function and generate output. The unit function iteration is sufficient, output response is satisfactorily gained. There are different kinds of exciting functions of which the most important are as follows (Fig. 2):

1. Threshold-logic functions
2. Hard-limit functions
3. Continuos functions (Sigmoid, ATAN)
4. Radial-basis or Gaussian function

The equation of Sigmoid function is:

$$Y = 1 / (1 + e^{-ax}) \tag{10}$$

where a = is constant; X= the sum of input exciting; Y= the sum of output excitation

Due to problems raised by indifferentiable functions, continuous functions have been presented, which are demonstrated by the sigmoid function. This function varies between two symmetric extremes. The two extremes are normally bonded between –1 and +1. For large values of X, these functions have positive derivatives.

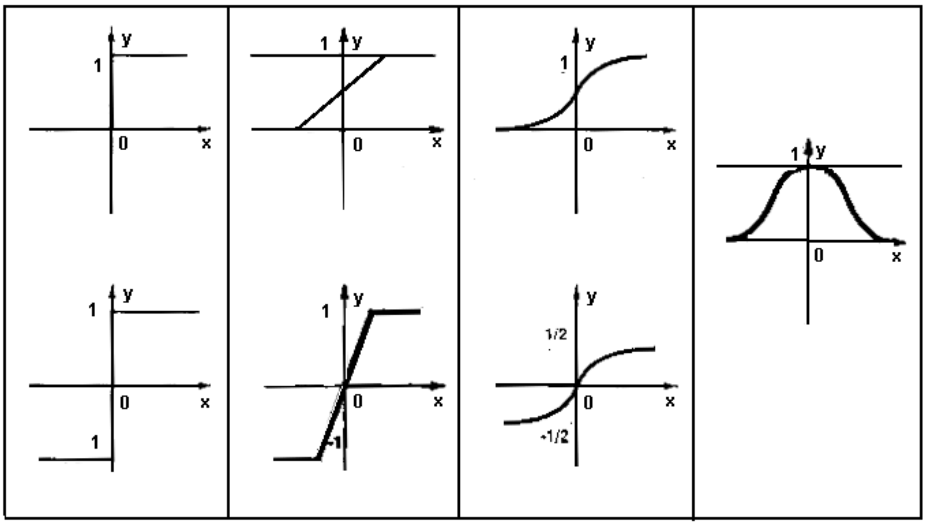

Fig. 2 Different kinds of exciting functions (Dayhoff 1990)

As initial and final values of weights, both before and after training, were comparable, the data of the input channel was scaled to match the first hidden layer range of the transfer function. For the sigmoid function, a scaling range between -0.8 and +0.8 was used.

In a back error propagating algorithm with a gradient descent rule, the learning rate is reduced as the error is propagating backward to the input channel, or in other words, the gradient gets attenuated by each layer in this reverse propagation. So it was preferred to use small learning rates for layers near the output channel, and increasing learning rates moving back toward the input channel. The standard steepest descent occurred when the learning rate is held constant through training, for high learning rates the algorithm may oscillate and become unstable, and for small learning rates the algorithm will converge. The adaptive learning ratse used in this study, kept learning step sizes as large as possible and made it responsive to the complexity of the local error surfaces. In the adaptive learning rates the weights and biases are corrected at each epoch using the correct learning rate. If the error in the current epoch is less than of the previous one, the learning rate is increased by about 5 to 10 percent, or is otherwise decreased. Also a momentum term was added in the modification of weights, but the momentum factor "e" remained unchanged during training:

$$\Delta w_{ij}^{n+1} = e\,(d_{pi} \cdot a_{pi}) + a\,\Delta w_{ij}^{n} \qquad (11)$$

Based on theoretical investigations of a three layer feed forward neural network, it shows that it is usually well suited to regression problems if enough hidden neurons are implemented, and it is also shown that for such problems more layers can improve the network performance. Generally the best architecture for the network depends on the problem presented to it. Except for purely linear networks, more neurons in the hidden layer result in a more powerful network. First the number of hidden layers was fixed to one and as the learning progressed, more and more neurons were added to the hidden layer until convergence occurred. For faster convergence rates, this scheme was developed for more hidden layers with equally increasing numbers of neurons in these layers, and finally the most suitable architecture was obtained. This is the case when the network does not converge for a reasonable number of neurons and a certain number of learning epochs or cycles under a predefined error gradient tolerance. This case is sometimes called trigger slope.

3 Analysis of water distribution system

3.1 Direct analysis

The effects of damages on hydraulic parameters (i.e. pressure, velocity and discharge) are divided into two categories. The first case is related to the damage that results in water leakage but water flow is not interrupted. The second case is related to the damage that lead the escape of all the water, causing the water flow to be disconnected. The first case takes place due to cracks in pipe bodies or valves being pushed in or pullet out. The second case occurs in pipes as breakage, segregation, step downs damage to joints, pulling out or the slipping out of valves occurs (Takada 2002).

3.1.1 Hydraulic models of flow in Leakage State

The rate of flow caused by leakage depends on the following factors: Water supply network pressures, Shape and diameter of holes or cracks, types, degree of aggregation, compaction of soil around the pipe and burial depth. One of the main factors of pressure is concerned with the leakage point. In all the water supply networks, some leakage is inherent. The America Water Work Association (AWWA 1994) consider leakage as 0.2 cubic meters per hour for each kilometer of pipe length in its calculations and design. Not surprisingly, the rate of leakage when an earthquake occurs suddenly increases.

Leakage discharge is obtained by Orifice rule (Eq. 12) (Streeter & wylie 1981).

$$Q_i = K_i\,(P_i - P_o)^{\,n} \qquad (12)$$

Where:$_i$ = leakage discharge (m^3/s), K_i = constant coefficient with respect to shape and hole diameter and P_i ,P_o= upstream pressure and downstream pressure of the hole (m)

Since the pressure in the downstream is always equal to zero, the equation above changes to the following:

$$Q_i = K_i P_i^{n} \qquad (13)$$

Based on Bernoulli's rule (Streeter & wylie 1981), since the velocity and discharge depend on the square root of height, the amount of 'n' will be 0.5 (n=0.5). Application of the Eq. (13) is very difficult. Determination of k depends on the shape and diameter of the hole. The pressure at the hole location is also unknown.

Eq. (14) is used for the calculation of leakage discharges. This equation denotes an experimental relationship and there is a big difference between it and the theoretical relationship. The advantage of Eq. (14) is the dependence on pipe pressure and the shape of the network.

$$Q_{ij} = L_{ij} \cdot (P_{ave})^{1.18} \cdot C \tag{14}$$

where Q_{ij} = Leakage along the pipe between nodes i and j (m^3/s); L_{ij} = Length of pipe in the distance of i and j (m);P_{ave} = Average pressure in the distance of i and j (m); and C = Constant coefficient.

In Eq. (14) the main factor is network pressure. The disadvantage of this relationship is that the pipe diameter is not included.

Finally Eq. (15) for various kinds of diameters is suggested.

$$Q_{ij} = L_{ij}.(P_{ave})^{1.18}.C_1.D_{ij}^{-1} \tag{15}$$

In this relationship the pipe diameter along ij segment is shown by D in meter.

3.1.2 Hydraulic model of flow in complete breakage state

The discharge of outlet flow caused by breakage and segregation of pipes, from a theoretical point of view, can be obtained from Eq. (16) (Streeter & wylie 1981):

$$Q = A_e.V_e \tag{16}$$

Where A_e is effective area and V_e is effective velocity of water.

4 Back analysis

When a breakage occurs in one of the network pipes, the tank shows an outlet discharge. The amount of outlet water discharges from the tank depends on the following factors (Streeter & wylie 1981):

1 Network pressure (the height of tank)
2 The diameter of the broken pipe
3 The distance of breakage point from the tank and
4 The height of the broken pipe with respect to the tank
5 The shape and form of the breakage
6 Soil and its compaction in the breakage point
7 Geometrical shape of network and the number of branches (the pipes among main nodes)
8 Technical characteristics and network equipment

When the network contains two tanks, each breakage affects the distance and diameter of the broken pipe, a known discharge from tank 1 and another from the tank 2 outflow. Each certain network is analysed by looking at all probable states of breakage. The break point is modeled as an imaginary tank with topographic height of the ground. Various probable breakage states are calculated as Eq. (17).

$$C_m^n = \frac{n!}{m! \cdot (n-m)!} \tag{17}$$

Where; C_m^n =Combination of m from n; n = Total number of pipes; m =The number of damaged pipe.

This massive number of analyses as well as data input number is impossible. By using ANN the rest can be estimated. In this study back propagation network (BPN) was applied.

4 Case study

A water distribution network with 2 tanks is presented here (Fig. 3).

Head of tanks: H= 45 m; Number of pipes = 7; Number of nodes = 6; Number of lopes = 2

Length and diameter of pipes 1 & 7 = 400m x 500mm

Length and diameter of pipes 2 & 6 = 400m x 400mm

Length and diameter of pipes 3 & 5 = 300m x 300mm

Length and diameter of pipe 4 = 300m x 250mm

Various probable breakage states are calculated. In total, 127 probable states of breakage occurrence exist in the example small network. This number of analyses as well as the data input would be difficult to undertake, even in this mentioned network. By using ANN, having had some random input data, the rest can be estimated by back propagation network (BPN).

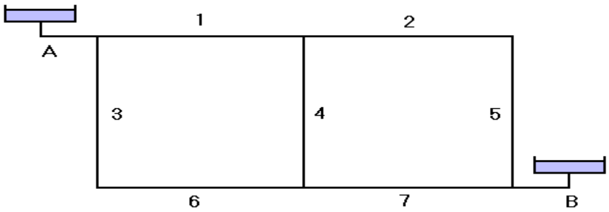

Fig. 3 A small water distribution network

4.1 Efficiency analysis

Due to the space consideration, some of the results are selected and showcased here. In order to perform the adaptive ascending algorithm, the parameters shown in table 1 were applied to the program and the number of hidden layers and hidden layer neurons were changed. As shown in Fig. 4 for 15 hidden neurons the networks with 1, 2 and 3 hidden layers were converged to the threshold limit within the existing epochs, and the first one gave the least root mean square (RMS) error. After each convergence the effect of over-training or over-fitting can be seen. In the next step for the selected hidden layer, a different number of hidden neurons were examined, which the results of some of them are presented in Fig. 5. The networks with 15 hidden neurons converged. The final network architecture shown in Fig. 6, and contained one hidden layer and fifteen hidden neurons in each.

Table 1 Assumed network parameters in the case study

Parameter	Value
Number of input neurons	2
Number of output neurons	7
Number of hidden layers	1
Number of hidden neurons	15
Activation function	Sigma

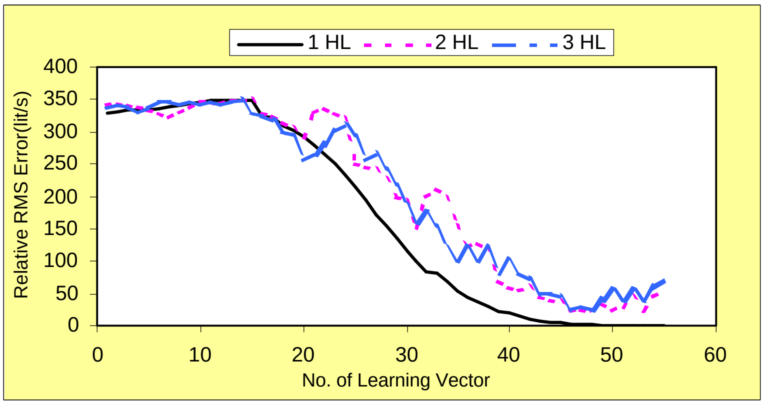

Fig. 4 Network error for different number of hidden layers

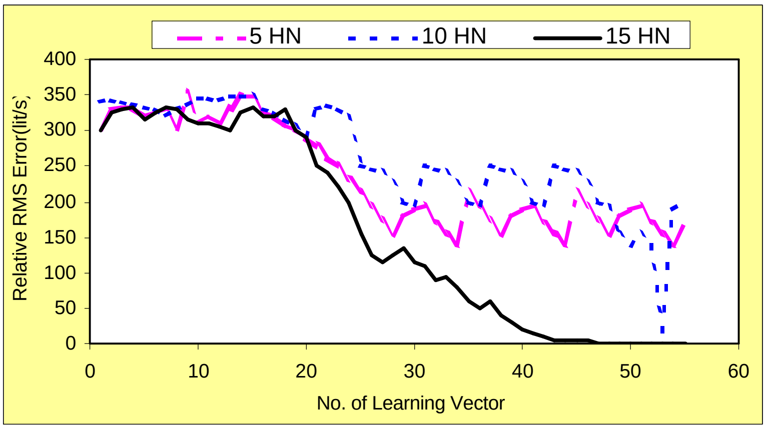

Fig. 5 Network error for different number of hidden layer neurons

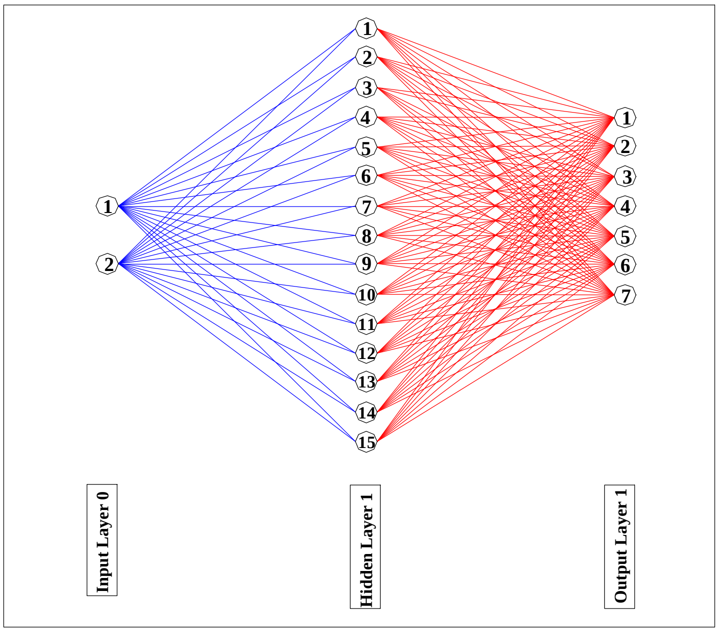

Fig. 6 Final network architecture purposed for the problem

5 Discussion and Conclusions

A general back error propagating perceptron with sigmoid transfer function was used to model the earthquake damage detection of water distribution systems. A direct analysis of a water network was implemented to create the data for a supervised learning procedure. The adaptive ascending algorithm with full connections were used to construct the best network architecture (Flood & Kartam 1994). The multiple random data generation and variable learning rate was used to speedup the learning process.

Parametric studies were also performed for different network parameters. It is noted that each problem with different geometric data, learning pattern and network architecture may be changed.

From the results presented in this study it is concluded that neural networks show great promise in becoming useful tools in practical analysis applications especially for problems with severe nonlinearities.

Other points and concluding remarks are as follows:

- The only method from back analysis method which is practical for determination of damage in water supply networks is the method of using outlet discharge of tanks.
- The program written by this method has a general state and is applicable for all kinds of water supply networks.
- Data base table for all kinds of networks based on geometrical shape, type, number and the characteristics of pipes and tanks by means of direct analysis is made up.

REFERENCES

American Water works Association, 1994. *Minimizing Earthquake Damage*, A Guide for Water Utilities.

Dayhoff, J.E. 1990. *Neural Network Architectures*, Van Nostrand Reinhold, New York.

Flood, I. Kartam, N. 1994. *Neural Networks in Civil Engineering*, ASCE, Vol, 2, No. 8, PP. 48-131, 1994.

Masri, S.F., Smyth, A.W., Chassiakos A.G., Caughey, T.K. & Hunter, N.F. 2000. Application of neural networks for detection of changes in nonlinear systems, *Journal of Engineering Mechanics* ASCE, Vol. 126, No. 7.

Streeter, L.V Wylie, E.B. 1981. *Fluied mechanics*, 7 th Edition, Mc Graw Hill, New York.

Takada, S. 2002. *Lifelines Earthquake Engineering*, Chapter 1 English version, under preparation.