

Selection of Proper Activation Functions in Back-propagation Neural Network algorithm for Single-Circuit Transmission Line

N. Suttisinthong, B. Seewirote, A. Ngaopitakkul, and C. Pothisarn, *Member, IAENG*

Abstract—This paper proposes an appropriate activation function for the fault classification decision algorithm. The decision algorithm based on the hybrid of discrete wavelet transform (DWT) and back-propagation neural network (BPNN) has been proposed to classify the fault type. The DWT is employed to decompose high frequency component of current signals. The maximum coefficient from the first scale at $\frac{1}{4}$ cycle of phase A, B, and C of post-fault current signals and zero sequence current obtained by the DWT have been used as an input variable in a decision algorithm. The activation functions in each hidden layer and output layer have been varied, and the results obtained from the decision algorithm have been investigated with the variation of fault inception angles, fault types, and fault locations. The results have illustrated that the use of Hyperbolic tangent sigmoid function in the first and the second layers with Linear function in the output layer is the most appropriate scheme for the transmission system.

Index Terms—Neural Networks, Discrete Wavelet Transform, Power Transformer

I. INTRODUCTION

ARTIFICIAL neural networks (ANNs) techniques have been proposed in some approaches in the literature to improve protective relay [1-9] due that these algorithms can give precise results. ANNs is characterized by (1) its pattern of connection between the neurons (called its architecture), (2) its method of determining the weights on the connections (called its training or learning algorithm), and (3) its activation function. Back-propagation neural network (BPNN) is a kind of neural networks, which is widely applied today owing to its effectiveness to solve almost all types of problem. Therefore, a decision algorithm, used for fault diagnosis in the power system to decrease complexity and duration of maintenance time, is required. Generally, the basic operation of an artificial neuron involves most units in neural networks transform their net input by using a scalar-to-scalar function called an "activation function", yielding a value called the unit's "activation". The activation function is sometimes called a

"transfer function". As a result, the activation function is a key factor in the artificial neural network structure. Back-propagation neural networks support a wide range of activation functions such as sigmoid function, linear function, and etc. The choice of activation function can change the behavior of the BPNN considerably. There is no theoretical reason for selecting a proper activation function. Hence, the objective of this paper is to investigate an appropriate activation function for the fault diagnosis decision algorithm. The activation functions in each hidden layers and output layer are varied, and the results obtained from the decision algorithm are studied.

The decision algorithm is a part of a transmission system scheme proposed in this paper. It is noted that the discrete wavelet transform is employed for extracting the high frequency component contained in the fault signals. The construction of the decision algorithm is detailed and implemented with various case studies based on Thailand electricity transmission and distribution systems.

II. SIMULATION

The ATP/EMTP is used to simulate fault signals at a sampling rate of 200 kHz. The fault types are chosen based on the Thailand's transmission system as shown in Fig. 1. Fault patterns in the simulations are performed with various changes in system parameters as follows:

- Fault types considered in this study are : single line to ground (SLG : AG, BG, CG), double-line to ground (DLG : ABG, BCG, CAG), line to line (L-L : AB, BC, CA) and three-phase fault (3-P : ABC).
- Fault locations are varied from 10% to 90%, with the increase of 10% of the transmission line length measured from the bus MM3.
- Inception angle on a voltage waveform is varied between 0° - 330° , with the increasing step of 30° . Phase A is used as a reference.
- Fault resistance is equal to 10Ω

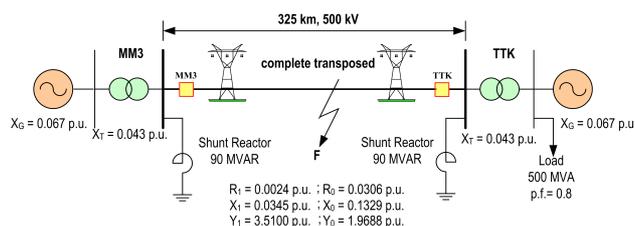


Fig. 1 The system used in simulation studies [1, 10].

Manuscript received January 15, 2014.

N. Suttisinthong, and B. Seewirote are with Department of Electrical Engineering, Faculty of Engineering, Thonburi University, Bangkok 10160, Thailand (e-mail: nuchtita@ideafield.co.th, srbancha7@gmail.com).

A. Ngaopitakkul, and C. Pothisarn are with Department of Electrical Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand (e-mail: knatthap@kmitl.ac.th, kpchaich@kmitl.ac.th).

The fault signals generated using ATP/EMTP are interfaced to MATLAB for the fault detection algorithm. Fault detection decision [1, 7] is processed using the positive sequence current signals. The Clark's transformation matrix is employed for calculating the positive sequence and zero sequence of currents. The mother wavelet daubechies4 (db4) [1, 4, 7, 9] is employed to decompose high frequency components from the signals. Coefficients obtained using the DWT of signals are then squared to clearly identify the abrupt change in the spectra. It is evident that the coefficients of high-frequency components are abruptly changed when a fault occurs, as shown in Fig. 2. The coefficients of scale 1 of DWT are used in the training processes for the neural networks in our case study.

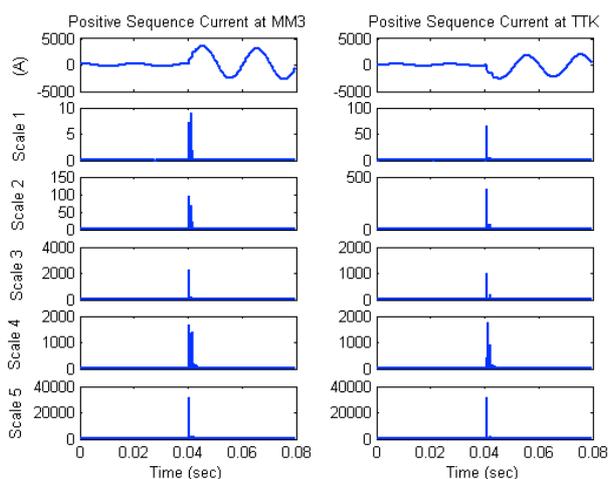


Fig. 2 Wavelet transform from scale 1 to 5 for the positive sequence of current signal.

III. DECISION ALGORITHM

After the fault detection process, the coefficients detail of scale 1, which is obtained using the DWT, is used for training and test processes of the back-propagation (BPNN). In this paper, a structure of a BPNN consists of three layers which are an input layer, two hidden layers, and an output layer as shown in Fig. 3. Each layer is connected with weights and biases. In addition, the activation function is a key factor in the BPNN structure. The choice of activation function can change the behavior of the BPNN considerably. Hence, the activation functions in each hidden

layers and output layer are varied as illustrated in Table 1 in order to select the best activation function. A training process was performed using neural network toolboxes in MATLAB which can be summarized in Fig. 4.

Table 1 Activation functions in all hidden layers and output layers for training neural networks

Activation function in		
first hidden layer	second hidden layer	output layer
Hyperbolic Tangent sigmoid function	Logistic sigmoid function	Linear function
		Logistic sigmoid
		Hyperbolic Tangent sigmoid
Logistic sigmoid function	Hyperbolic Tangent sigmoid function	Linear function
		Logistic sigmoid
		Hyperbolic Tangent sigmoid
Logistic sigmoid function	Logistic sigmoid function	Linear function
		Logistic sigmoid
		Hyperbolic Tangent sigmoid
Hyperbolic Tangent sigmoid function	Hyperbolic Tangent sigmoid function	Linear function
		Logistic sigmoid
		Hyperbolic Tangent sigmoid

By observing Fig. 4, before the fault classification decision algorithm process, a structure of the BPNN consists of 4 inputs. The maximum coefficients detail (phase A, B, C and zero sequence of post-fault current) of DWT at the first peak time that can detect fault, is performed as input variables. The output variables of the BPNN for identifying fault types are designated as either 0 or 1, corresponding to phases A, B, C and ground (G). If each output value of BPNN is less than 0.5, no fault occurs on each phase; conversely, if this output value of BPNN is more than 0.5, a fault does occur. The output variables of the BPNN for locating the fault along the transformer windings are designated as values range of 0.1 to 0.9, corresponding to length of the winding that fault occurs. In addition, before training process, it is important to prepare the data before analysis because the obtained data will be clearly analyzed for the pattern between input pattern and output pattern. Input data sets define value range of 0 to 1 by normalization.

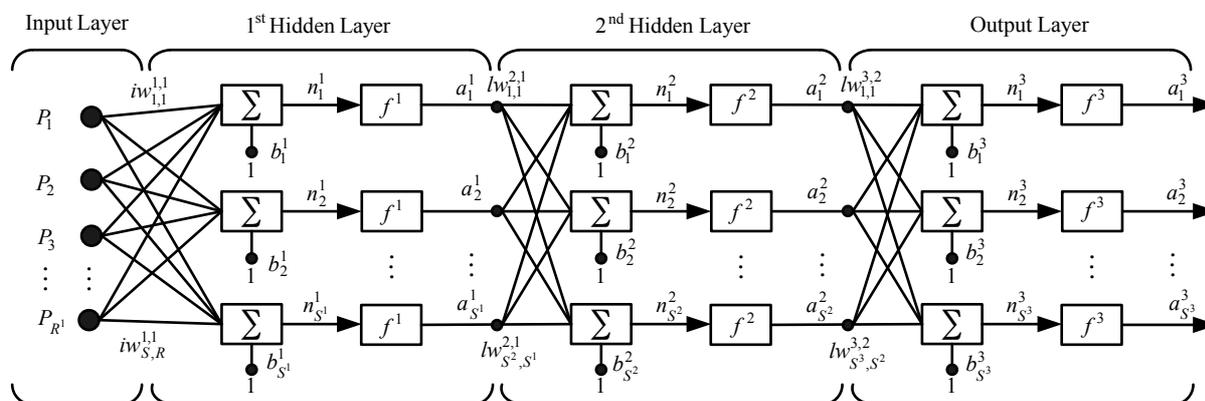


Fig. 3. Structure of BPNN.

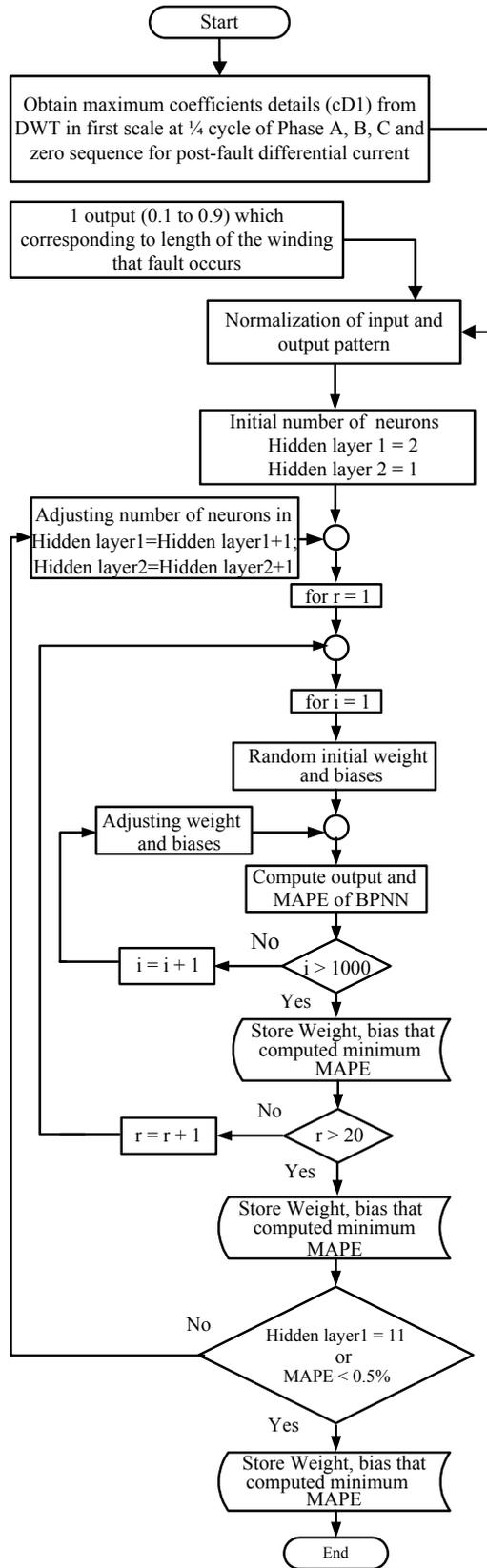


Fig. 4. Flowchart for the training process

A training process for back-propagation neural network can be divided into three parts as follows [1]:

1. The feedforward input pattern, which has a propagation of data from the input layer to the hidden layer and finally to the output layer for calculating responses from input patterns illustrated in Equations 1 and 2.

$$a^2 = f^2(lw^{2,1} * f^1(iw^{1,1} * p + b^1) + b^2), \quad (1)$$

$$o / p_{ANN} = f^3(lw^{3,2} * a^2 + b^3). \quad (2)$$

where,

p is the input vector of ANNs

$iw^{1,1}$ is the weights between input and the first hidden layer

$lw^{2,1}$ is the weights between the first and the second hidden layers

$lw^{3,2}$ is the weights between the second hidden layer and output layers

b^1, b^2 are the bias in the first and the second hidden layers respectively

b^3 is the bias in output layers

f^1, f^2 are the activation functions (Hyperbolic tangent sigmoid function: tanh)

f^3 is the activation function (Linear function)

2. The back-propagation for the associated error between outputs of neural networks and target outputs; the error is fed to all neurons in the next lower layer, and also used to an adjustment of weights and bias.

3. The adjustment of the weights and bias by Levenberg-Marquardt (trainlm). This process is aimed at trying to match between the calculated outputs and the target outputs. Mean absolute percentage error (MAPE) as an index for efficiency determination of the back-propagation neural networks is computed by using Equation 3.

$$MAPE = \frac{1}{n} * \sum_{i=1}^n \left| \frac{o / p_{ANNi} - o / p_{TARGETi}}{o / p_{TARGETi}} \right| * 100\% \quad (3)$$

where, n is the number of test sets.

During training process, the weight and biases are adjusted by Levenberg-Marquardt (trainlm), and there are 20,000 iterations in order to compute the best value of MAPE. The number of neurons in both hidden layers is increased before repeating the cycle of the training process. The training procedure is stopped when reaching the final number of neurons for the first hidden layer or the MAPE of test sets is less than 0.5%.

After training process, the decision algorithm is employed to identify the fault types in the transmission line. Case studies were performed with various fault types at each location along the transmission lines and varying fault inception angles and locations along each transmission line, as shown in Fig. 5 and Fig. 6. By observing Fig. 5, the comparison of average accuracy among the type of fault for various lengths of the transmission lines that fault occurs is presented. The comparison between an average accuracy in fault identifications obtained from the various activation functions is shown in Fig. 6. From Fig. 6, it is shown that when the case studies are tested with various locations at each transmission line, the accuracy of fault type from the BPNN algorithm is highly satisfactory. In addition, the proposed algorithm Hyperbolic tangent – Hyperbolic tangent – Linear as activation function in each layer can give a better performance in predicting the fault types.

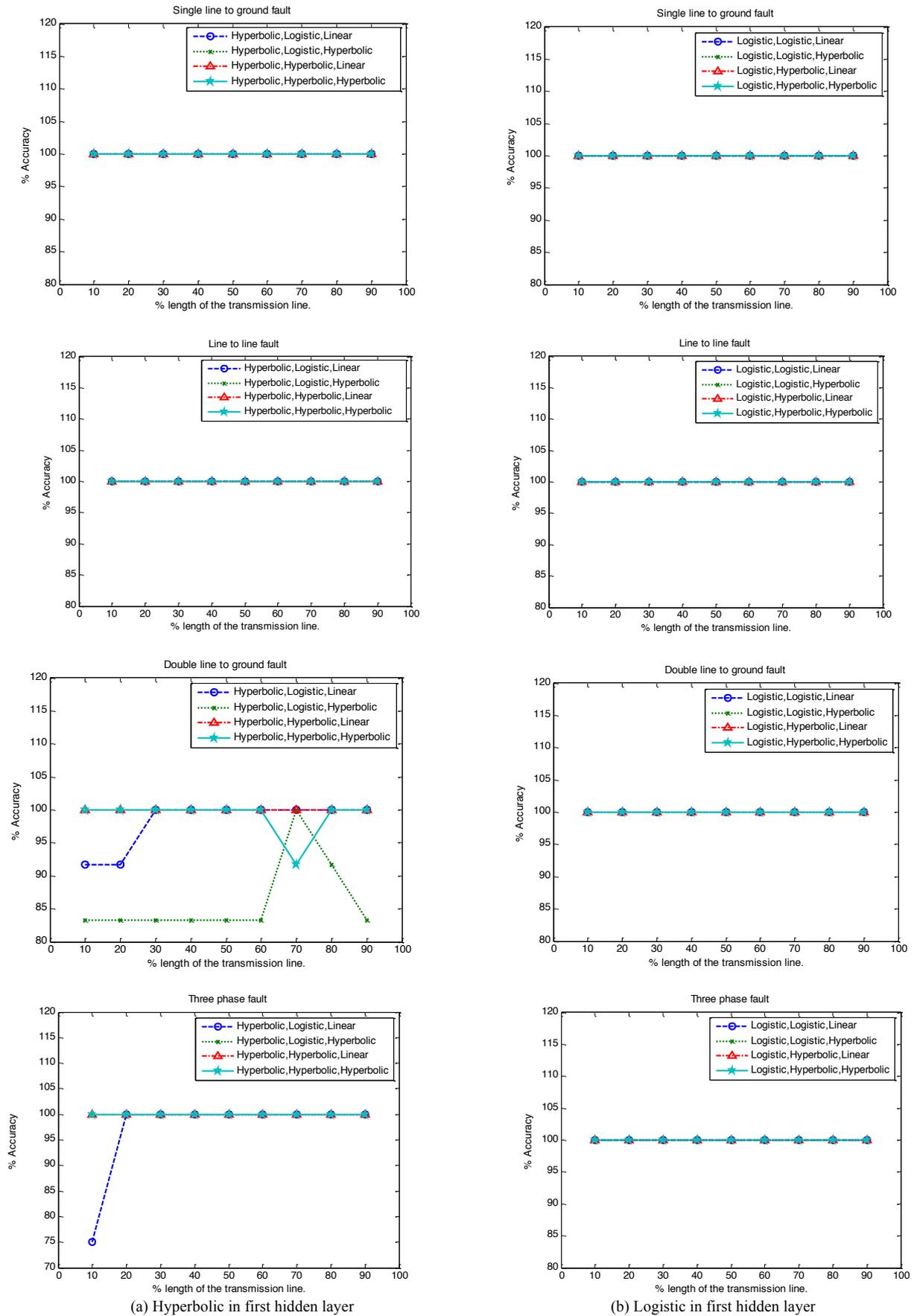


Fig. 5. Comparison of average accuracy when identifying the fault type at various lengths of the transmission line among various fault types.

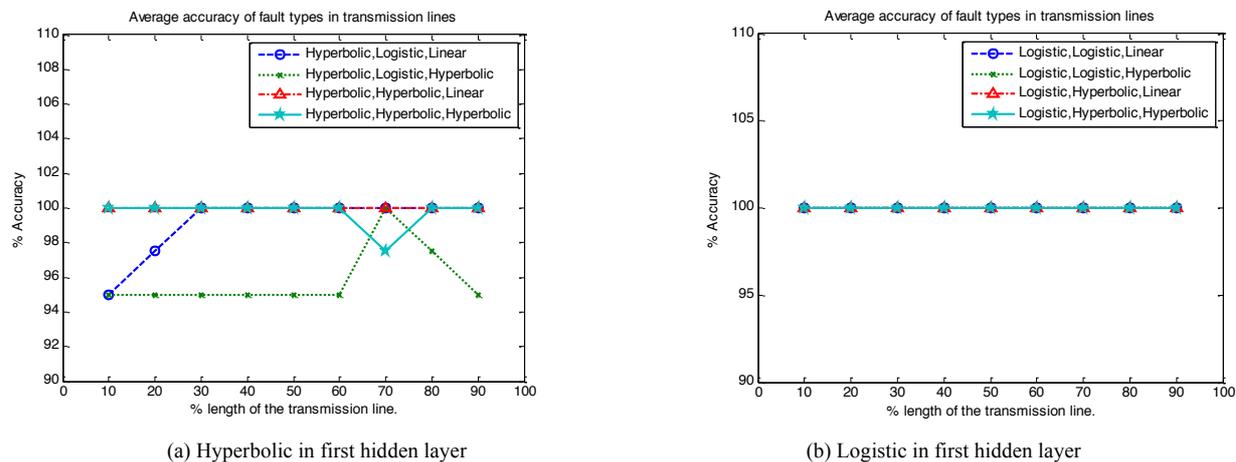


Fig. 6. Comparison of average accuracy when identifying the fault type at various lengths of the transmission line among various activation functions.

IV. CONCLUSION

In this paper, study of an appropriate activation function for the decision algorithm used in the transmission system has been discussed. The fault classification decision algorithm can be classified the fault type and fault location using the back-propagation neural networks (BPNN). The maximum coefficient from the first scale at $\frac{1}{4}$ cycle of phase A, B, and C of post-fault current signals and zero sequence current obtained by the DWT have been used as an input variable in a decision algorithm. The activation functions in each hidden layer and output layer have been varied, and the results obtained from the decision algorithm have been investigated with the variation of fault inception angles, fault types, and fault locations. The results have illustrated that the use of Hyperbolic tangent sigmoid function in the first and the second layers with Linear function in the output layer is the most appropriate scheme for the transmission system.

ACKNOWLEDGMENT

This work is partially supported by the King Mongkut's Institute of Technology Ladkrabang Research fund. The authors would like to thank for this financial support. The authors would like also to gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] P. Chiradeja and A. Ngaopitakkul, "Identification of Fault Types for Single Circuit Transmission Line using Discrete Wavelet transform and Artificial Neural Networks," *In Proceedings of 2009*

International Conference on Electrical Engineering (ICEE2009), pp. 1520-1525, 2009.

[2] N. Perera, A.D. Rajapakse, "Recognition of Fault Transients Using a Probabilistic Neural-Network Classifier," *IEEE Transactions on Power Delivery*, Vol. 26, No. 1, pp. 2011, 410 – 419.

[3] D. Barbosa, D.V. Coury, and M. Oleskovicz, "New approach for power transformer protection based on intelligent hybrid systems," *IET Generation, Transmission & Distribution*, Vol. 6, No. 10, 2012, pp. 1009-1018.

[4] A. Ngaopitakkul, C. Jettanasen, J. Klomjit, C. Pothisarn, B. Seewirote, and N. Suttisinthong, "Application of back-propagation neural network for transformer differential protection schemes part 1 discrimination between external short circuit and internal winding fault," *Joint 6th International Conference on Soft Computing and Intelligent System (SCIS) and 13th International Symposium on Advanced Intelligent Systems*, 2012, pp. 1493-1498.

[5] Yang Cunxiang, Bao Hao, "The fault diagnosis of transformer Based on the SOM neural network current," *Proceedings of the 5th Conference on Measuring Technology and Mechatronics Automation*, 2013, pp. 1178-1180.

[6] S. Seyedtabaai, "Improvement in the performance of neural network-based power transmission line fault classifiers," *IET – Generation, Transmission and Distribution*, Vol. 6, No. 8, pp. 731-737, 2012.

[7] A. Ngaopitakkul and C. Jettanasen, "Combination of Discrete Wavelet Transform and Probabilistic Neural Network Algorithm for Detecting Fault Location on Transmission System," *International Journal of Innovative Computing, Information and Control*, Vol. 7, No. 4, pp. 1861-1874, 2011.

[8] Y.Zhang, X.Ding, Y.Liu and P.J. Griffin, "An artificial neural network approach to transformer fault diagnosis", *IEEE Trans. Power Delivery*, Vol. 11, No. 4, pp. 1836-1841, 1996.

[9] A. Ngaopitakkul and A. Kunakorn, "Internal Fault Classification in Transformer Windings using Combination of Discrete Wavelet Transforms and Back-propagation Neural Networks," *International Journal of Control, Automation, and Systems (IJCAS)*, Vol. 4, No. 3, pp. 365-371, June 2006.

[10] "Switching and Transmission Line Diagram", *Electricity Generation Authorisation Thailand (EGAT)*.