World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:2, No:6, 2008

Implementation of Neural Network Based Electricity Load Forecasting

Myint Myint Yi, Khin Sandar Linn, and Marlar Kyaw

***Abstract*—**This paper proposed a novel model for short term load forecast (STLF) in the electricity market. The prior electricity demand data are treated as time series. The model is composed of several neural networks whose data are processed using a wavelet technique. The model is created in the form of a simulation program written with MATLAB. The load data are treated as time series data. They are decomposed into several wavelet coefficient series using the wavelet transform technique known as Non-decimated Wavelet Transform (NWT). The reason for using this technique is the belief in the possibility of extracting hidden patterns from the time series data. The wavelet coefficient series are used to train the neural networks (NNs) and used as the inputs to the NNs for electricity load prediction. The Scale Conjugate Gradient (SCG) algorithm is used as the learning algorithm for the NNs. To get the final forecast data, the outputs from the NNs are recombined using the same wavelet technique. The model was evaluated with the electricity load data of Electronic Engineering Department in Mandalay Technological University in Myanmar. The simulation results showed that the model was capable of producing a reasonable forecasting accuracy in STLF.

***Keywords*—**Neural network, Load forecast, Time series, wavelet transform.

I. INTRODUCTION

IN the forecasting of electricity load, short term load forecasting (STLF) has become popular. STLF includes forecasting of electricity load of a few minutes, hours, or days ahead. The purpose of STLF is to predict the future electricity load base on the historical load data. There are two different forecasting models- the traditional models and the modern techniques. Traditional forecast models employ times series and regression analysis through the use of statistical models such as peak load models and load shape models. The load shape modes relay on time series analysis techniques. The autoregressive moving average (AVMA) model is one of the most popular dynamic load shape models.

Artificial intelligence methods for forecasting give better performance in modeling of the time series. In this paper, we study neural network models combined with wavelet decomposition. This approximates a time-series at different levels of resolution using multiresolution decomposition. Wavelet recombination technique is employed to reconstruct data after wavelet decomposition. With the help of this technique, a time series can be predicted as an additive combination of the wavelet coefficients at different resolution levels. The date to predict electricity demand using neural network is obtained from Electronic Department of MTU in Myanmar.

The proposed model is trained with back-propagation algorithm using multilayer perceptron (MLPs). This model is trained with supervised learning algorithm. The model adopts a learning algorithm which is known as the perceptron convergence algorithm (PCA). The PCA has two classifications of learning algorithms such as supervised and unsupervised learning algorithms. Supervised learning is presented with both the inputs as well as the expected output. Unsupervised learning is presented only inputs but not outputs. Neural network can be classifications according to the type of input data or the type of algorithm that is implemented. Neural networks which take in binary inputs are usually associated with logic functions. It can be solved logical functions with inputs taking values of 1 or 0.since the perceptron is capable of solving Boolean functions. Neural networks that take in continuous type of inputs or time series data can be presented with any range of values for both its input and output vectors. This type of neural network is used for application such as pattern classification and function approximation. Multi-layer feed-forward neural network is used if a supervised learning algorithm is adopted. The multi-layer feed-forward neural network is also a popular architecture which has been proven to give satisfactory results for time-series predictions [1] [2]. MLPs are feed-forward neural network commonly trained in a supervised algorithm with the error back-propagation (BP) algorithm [3] [4].

II. WAVELET DECOMPOSITION IN FORECASTING

A. Wavelet Decomposition

Wavelet decomposition provides a signal to break down into many lower resolution components. This is called the wavelet decomposition tree.

Wavelet decomposition tree can yield a signal to get valuable information. The decomposition can continue only until the individual details consist of a single sample or pixel. We can select a suitable number of levels based on the nature of the signal [6].

Manuscript received March 23, 2008. This work was supported in part by the Ministry of Science and Technology, Union of Myanmar.

Myint Myint Yi is with the Mandalay Technological University Myanmar, (phone: 095-02-88702; fax: 095-02-88702; e-mail: m.myintyi@ gmail.com).

Khin Sandar Linn is with the Mandalay Technological University, Myanmar (phone: 095-02-88702; fax: 095-02-88702; e-mail: khinsandarlinn.mm@ gmail.com).

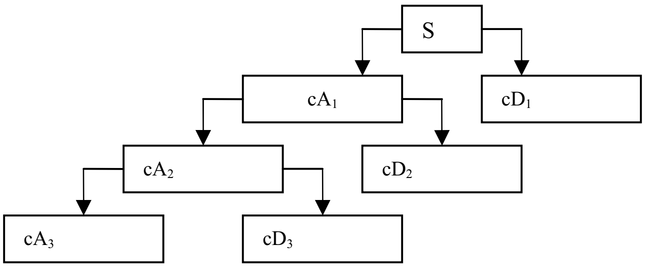

Fig. 1 Wavelet decomposition
(Figure obtained from MATLAB help file)

B. Wavelet Reconstruction

Wavelet reconstruction provides the components to assemble back into the original signal without loss of information. This process is called reconstruction.

The mathematical manipulations that effects synthesis is called the inverse discrete wavelet transform (IDWT).

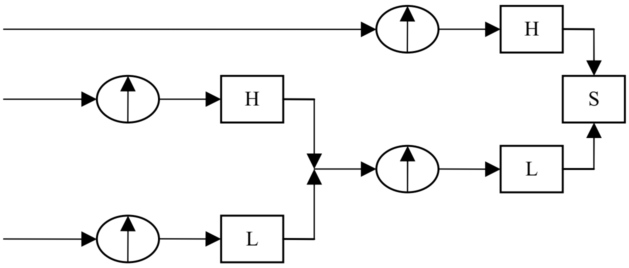

Fig. 2 Wavelet reconstruction

Wavelet decomposition involves filtering and down sampling. The wavelet reconstruction process consists of up sampling and filtering. Up sampling is the process of lengthening a signal component by inserting zeros between samples.

The down sampling of the signal components performed during the decomposition. During the decomposition, it is necessary to choose the correct filters. There are two types of filters to decompose or reconstruct the signal. The low and highpass decomposition filters (L and H), together with their associated reconstruction filters (L and H), form a system [6].

C. Reconstructing Approximations and Detail

It is possible to reconstruct the original signal from the coefficients of the approximations and details. It is also possible to reconstruct the approximations and details themselves from their coefficients vectors. As an example, we can reconstruct the first-level approximation A1 from the coefficient vector cA1.

We reconstruct the original signal instead of combining it with the level-one detail cD1; we can feed in a vector of zeros in place of the detail coefficients vector.

The process gives a reconstructed approximation A1, which has the same length as the original signal s which is a real approximation. We can also reconstruct the first-level detail D1 by using the analogous process.

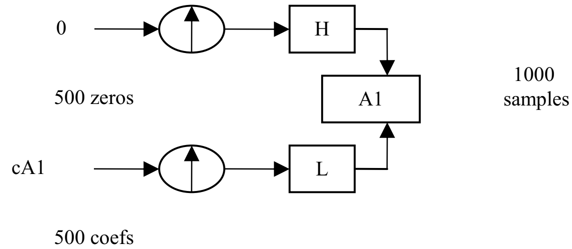

Fig. 3 Wavelet reconstruction from approximation and detail vectors

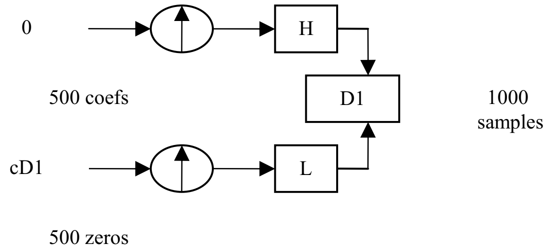

Fig. 4 Wavelet reconstruction for detail vector

The reconstructed details and approximations are true parameters of the original signal.

$$A_1 + D_1 = S \tag{1}$$

The coefficient vectors cA1 and cD1 are produced by downsampling and are only half the length of the original signal. Therefore, they cannot directly combine to reproduce the signal. It is necessary to reconstruct the approximations and details before they combine with each other.

There are several ways to reconstruct the original signal from the coefficients. They are shown in the following figure [6].

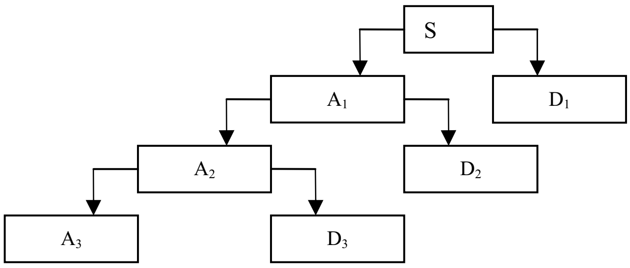

Fig. 5 Wavelet reconstruction from approximation and detail vectors

D. Neural Network Model for Forecasting of Time Series Signal

Most input signals are time series in their raw format. Most signals are measured as functions of time. The forecast model for electricity load is shown in figure. In this figure the raw data input is a noisy time series pattern, which can affect the accuracy of the forecast model's prediction. To produce the best quality of the raw input signal for time series forecast, the neural network model is used with multiple resolutions.

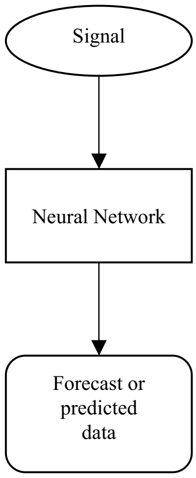

Fig. 6 Neural network model for forecasting

This multiple resolution is based on wavelet transform. The wavelet transform can be divided into three steps. In the first step, the input time series raw data is pre-processed using wavelet decomposition. Then each decomposed scale put into one neural network as its input and the respective neural networks (NNs) take in their allocated scale for training or for prediction in the second step. In the third step the decomposed scales at the output of the NNs are recombined to obtain the required prediction [5].

Wavelet transform (WT) is a scalable technique. The required size allows the use of long time intervals when low frequency information is required. It allows the use of short time intervals when high frequency information is required. There are three different wavelet transforms. They are continuous wavelet transform, discrete wavelet transform and non-decimated wavelet transform.

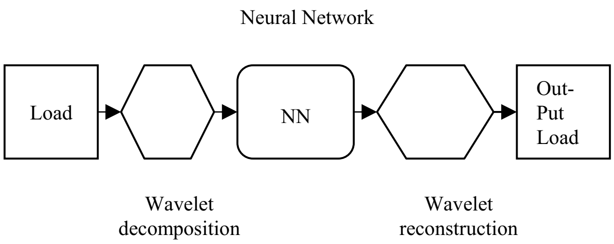

Fig. 7 Neural network model for time series forecasting

Continuous wavelet transform (CWT) is the sum of over all time of the signal multiplied by scaled, shifted versions of the wavelet function. Continuous wavelet transform is a set of continuous variables in nature and can create difficulties calculation of coefficients. Discrete wavelet transform (DWT) can produce coefficients of fine scales for capturing high frequency information. The non-decimated wavelet transform (NWT) is known as the undecimated or stationary wavelet transform. Non-decimated wavelet transform produces equal-length wavelet coefficients for each resolution levels. In this forecasting model, non-decimated wavelet transform is used [5].

III. Back-Propagation Algorithm

In the modern world today, many NN applications require the networks to solve diverse problems or to perform complex computations. Unfortunately, the algorithms that have been discussed so far are all formulated for single-layer NNs. To handle diverse problems or complex computations, the multi-layer perceptron (MLP) networks are introduced. Fig. 8 illustrates an example of a MLP with two hidden layers.

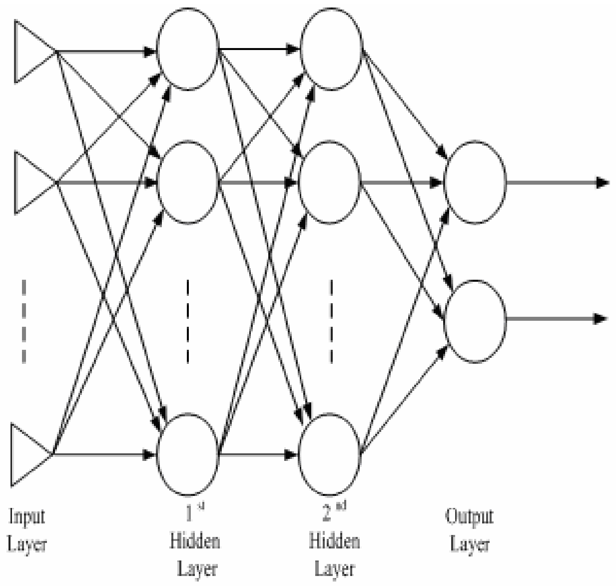

Fig. 8 Multi-layer perceptron with two hidden layer

MLPs are feed-forward NNs commonly trained in a supervised fashion with the error back-propagation (BP) algorithm [3] [4]. The BP algorithm is the generalization of the W-H learning algorithm to multi-layer networks and nonlinear differentiable transfer functions. The algorithm consists of a forward pass and a backward pass.

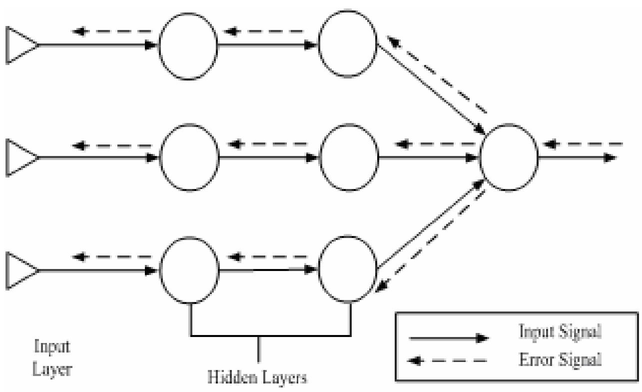

Fig. 9 Flow of signal with the black-propagation algorithm

It is applied by first passing an input signal forward through the network in the direction towards the output until a set of the actual response is obtained from the network. Fig. 9 illustrates the flow of the forward and the backward signals.

The error signal is then generated based on the difference between the actual and the target response.

Finally the generated error signal is passed backwards through the hidden layers, in the direction towards the inputs. During the forward pass, the synaptic weights of the network are fixed. It is during the backward pass that the synaptic weights are adjusted to adapt the network in producing desired outputs.

The adaptation of a BP MLP network with an arbitrary number of layers can be mathematically expressed as

$$\Delta w = \sum_{p} \delta_{layer} \times v_{input} \tag{2}$$

where η is the learning rate, V_{input} is the input vector appearing at the input terminal of the weight, and δ_{layer} is a term associated with the error function which is layer dependent.

The basic BP algorithm is a gradient descent algorithm which adjusts the network weights along the steepest descent direction of the error function (that is, the direction in which the error function decreases most rapidly; negative of the gradient).

Since the BP algorithm is widely employed with supervised MLP networks, it is not surprising that many variants of it are known to exist. These variants are commonly known as the Conjugate Gradient Algorithms (CGA). Table III is a summary of various popular CGAs obtained from the MATLAB help file

Back propagation is the basis for training a supervised neural network. The data used as inputs is transmitted through the network, layer by layer, and a set of outputs is obtained. During this forward pass the weights of the network are set. The obtained outputs are compared with the desired outputs values and, as a backward pass; the difference between desired outputs and calculated outputs (error) is used to adjust the weights of the net in order to reduce the level of the error.

This is an iterative process, which continues until an acceptable level of errors will be obtained. Each time the network processes the whole set of data (both a forward and a backward pass), is called an epoch. The network is in this way trained and the error is reduced by every epoch until an acceptable level of error will be gained. This method is called error back-propagation training.

IV. Methodology

The proposed model is tested with the electricity load data for one month in Electronic Engineering Department in Mandalay Technological University, Myanmar. The simulation results are obtained from the program written with MATLAB software. In this model, we can observe that the signal in time series can be decomposed into many scales by using wavelet transform. We can also observe back propagation algorithm to get error signal and various learning algorithms such as supervised or unsupervised learning algorithms.

V. Preliminary Forecast Model

The preliminary forecast model has three main stages. This is illustrated in Fig. 10.

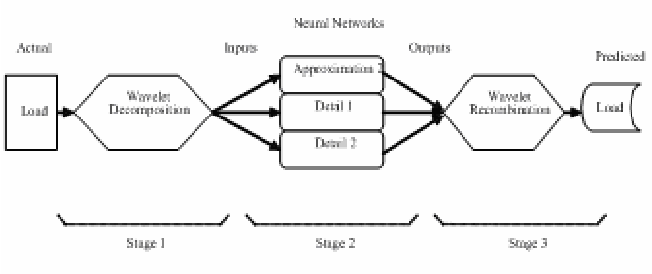

Fig. 10 Three stages of the preliminary forecast model

As depicted in figure, wavelet techniques are implemented in the first and the last stages. The actual time-series (electricity load and temperature data) are first decomposed into a number of wavelet coefficient signals and one approximation signal. The decomposed signals are then fed into the NNs at the second stage to predict the future time-series patterns for each of the signals. Finally, the predicted signals are recombined in the last stage to form the final predicted load time series.

A. Stage 1: Pre Signal Processing

For pre signal processing, historical data is fed to the model as time-series signals. Non-decimated Wavelet Transform (NWT) is used as the pre signal processor in the model and depending on the selected resolution level, the respective time-series signals are decomposed into a number of wavelet coefficients. In other words, if the resolution level is defined as n, after decomposing the signal, there will be one approximation coefficient series as well as n number of detail coefficient series. These decomposed coefficients are then normalized and fed as inputs to the signal predictor (NNs) for either training or forecasting use. The decomposition process is illustrated in Fig. 11.

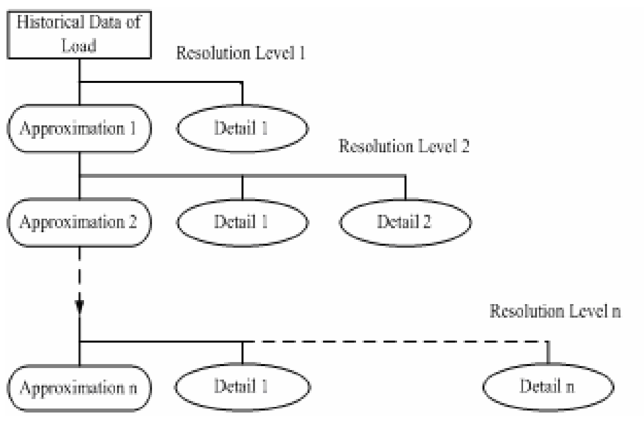

Fig. 11 Wavelet decomposition process

The most suitable resolution level is identified based on the smoothness of the approximation signal at that level (i.e. having all the high frequency components removed). The desired approximation signal should depict a general pattern of its original. A higher resolution level would produce a smoother approximation signal. The details on determining this resolution level will be discussed in the simulation section.

B. Stage 2: Signal Prediction

NNs are used for signal prediction in the forecast model. The number of NNs needed for the model is determined by the number of wavelet coefficient signals at the output of the pre-processor. For each wavelet coefficient signal (including the approximation), one NN is required to perform the corresponding prediction.

C. Stage 3: Post Signal Processing

For post signal processing, the same wavelet technique (NWT) and resolution level are used. In this stage, the outputs from the signal predictor (NNs) are combined to form the final predicted output. This is achieved by summing all the predicted wavelet coefficients. Fig. 9 illustrates the recombination process.

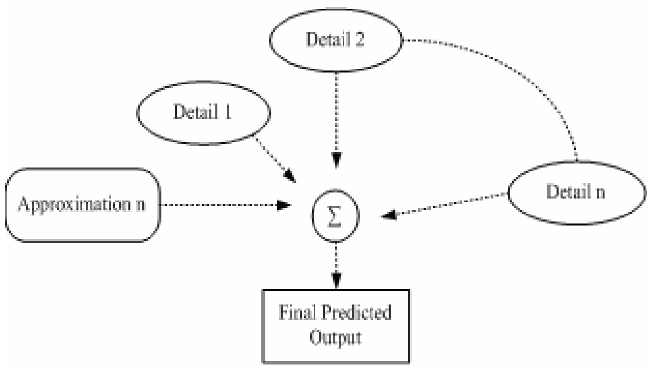

Fig. 12 Wavelet recombination process

VI. Simulation

As mentioned above, the best resolution level is tested. The results are as shown in Fig. 13. In Fig. 13, we can observe that the approximation signal at resolution level two is significantly smooth to represent a general pattern of the original signal. Therefore, resolution level two is chosen for both pre and post processing stages.

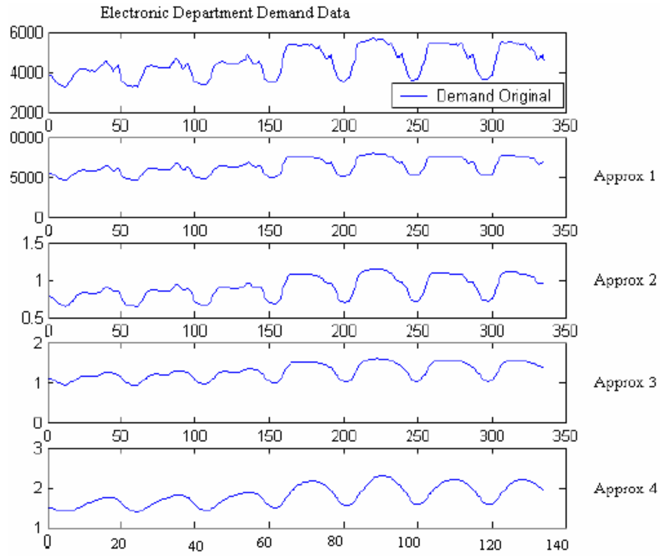

Fig. 13 Comparison of Different Resolution Levels

VII. Conclusion

This paper proposed a STLF model with a high forecasting accuracy. The NWT has been successfully implemented in the model. The implementation of NWT has reasonably enhanced the learning capability of the NNs in the model, thus minimizing their training frequencies as shown in the simulations. In summary, the inclusion of load data (as input variable) and the use of NWT (as the data processing tool) for the proposed STLF model have been a success.

Acknowledgment

The author is deeply grateful to Dr.Yin Mon Myint, Associate Professor and Head, Department of Electronic Engineering, Mandalay Technological University, for her willingness to share her ideas and helpful suggestions.

The author also wishes to thank Dr.Maung Maung Latt, Professor, Department of Electrical and Electronic Research Center, Mandalay Technological University, for his encouragement, help, support and guidance.

The author is extremely grateful thanks to U Clement Saldanha, Associate Professor, Department of Electrical and Electronic Research Center, Mandalay Technological University, for his skillful guidance and encouragement.

The author wishes to acknowledge especially her gratitude to Daw Marlar Kyaw, Lecturer, Department of Electronic Engineering and Information Technology , Mandalay Technological University, for her valuable advices, guidance, encouragement thought-out the course of thesis.

Many thanks go to all teachers whose assistance was vital for research. Finally, the author would like to thank her parents, elder brother and sisters whose love never ceases to amaze her. They have taught her many valuable lessons in life and have always given their best advice.

REFERENCES

[1] H.S. Hippert et al, "Neural Networks for Short-Term Load Forecasting: A Review and Evaluation," *IEEE Trans.on Power Systems,* Vol.16, No.1, Feb. 2001.

[2] W.R. Foster et al, "Neural Network forecasting of short, noisy time series," *Computers Chem. Engng*., Vol. 16, No.4, 1992.

[3] S. Haykin, "Neural Networks," Macmillan College Publishing Company, Inc.1994

[4] N.K. Bose and P.Laing, "Neural Network Fundamentals with Graphs, Algorithms, and Applications," McGraw- Hill Electrical and Computer Engineering Series, McGraw-Hill, Inc.1996.

[5] K.L.Toe, "Neural Network in Electricity Load Forecasting with Wavelet Decomposition and Temperature Dependency," Thesis of Electrical Engineering, University of Queensland, August.2001.

[6] "Wavelet Toolbox," MATLAB Software.

[7] H. Maier, 'A Review of Artificial Neural Networks," Department of Civil and Enviromental Engineering, University of Adelaide, Research Report No.R[3], August,1995.

[8] G. Chicoo et al, "Laod pattern clustering for short-term load forecasting of anomalous days," *IEEE Porto PowerTech 2001 Conferences, September,2001.*

[9] B.L.Zhang and Z.Y.Dong, An Adaptive Neural-Wavelet Model for Short Term Load Forecasting," Electrical Power Systems Research 59(2001).

[10] B.N. Tran et al, "Wavelets," *J.Webster(ed)* Wiley Encyclopedia of Electrical and Electronics Engineering, John Wiley & Sons, Inc, 1999.

[11] A. Louis et al, "Wavelets Theorey and Applications," John Wiley & Sons Ltd, England,1997.

Myint Myint Yi was born in 1978, November 27. Graduated in 10rd March 2002 with B.E (Electronic Engineering) and finished Master degree on July, 2003 with M.E (Electronic Engineering). At present, she is a Ph.D candidate of Electronic Engineering Department, MTU, and Mandalay, Myanmar.

She served as an Assistance Lecture at Hpa-An TU from September, 2003 to January, 2004 and then she was promoted to the Lecturer of Hpa-an Technological University, Department of Technical and Vocational Education, Myanmar. Now, she is making her PhD research at Mandalay Technological University (MTU), Myanmar with the title of "Design and Construction of Digital Printer Interface for Rate Meter'.

Khin Sandar Linn was born in 1982, March 17. Graduated in 3rd November, 2004 with B.E (Electronic Engineering) and finished Master degree on March, 2006 with M.E (Electronic Engineering). At present, she is a Ph.D candidate of Electronic Engineering Department, MTU, and Mandalay, Myanmar. Now, she is making her PhD research at Mandalay Technological University (MTU), Myanmar with the title of "Implementation of Leaf Identification System using Neural Networks".