Hybrid and Integrated Intelligent System for Load Demand Prediction

B.Islam, Z. Baharudin*, Q.Raza, P. Nallagownden
Department of Electrical & Electronics Engineering
Universiti Teknologi PETRONAS
31750 Tronoh, Preak, Malaysia
*zuhairb@petronas.com.my

***Abstract*—Artificial neural networks (ANN) are receiving a lot of attention because of their nonlinear mapping ability in the field of short term load forecast (STLF). ANN based STLF model commonly use back propagation algorithm, that may not converge properly, that affects the forecast accuracy. A hybrid approach, based on artificial neural network (ANN) and genetic algorithm (GA) that combines the advantages of each technique is proposed in this research. Genetic algorithm is implemented for the optimization of the architecture of feedforward neural network and selection of its initial weight values. Error back propagation algorithm for the training of the optimized neural network will be implemented. The second stage of this research is related with the complete training of the neural network based on genetic algorithm, using genetic manipulation of chromosomes. The results show that this approach produced better output in terms of enhanced forecast accuracy.**

***Keywords*—Back-propagation, Genetic Algorithm, Multi layer perceptron neural network .**

I. Introduction

Load forecast is a prediction of future load, which plays very important role in the energy management system [1] to provide a better planning for the power system. Most of the operational decisions such as economic scheduling of generating capacity, schedules of fuel purchases, infrastructure development and system security assessment rely on these predictions [1, 3-5].

Typically, load forecasting is classified according to planning duration, as the following [4, 6-8].

- Short term forecast with a lead time of up to a few hours ahead is emphasized for the economic and day to day operational decisions of systems.
- Medium term forecast with a lead time of few weeks to one year period is beneficial for the scheduling of fuel supplies and maintenance operations.
- Long term forecast with a lead time of up to ten years is efficacious for planning strategies of the power systems.

Short term load forecast (STLF) is vital for planning the daily operations of an electric power system [8-10]. Precise and authentic load demand prediction on hourly basis from one day to a week ahead assists the system operators to execute multiple tasks such as economic management of generating capacity and tabulating of fuel purchases. Particularly, peak demand forecast is of vital importance because the generating capacity of an electric utility should comply with this requirement [9, 11].

Excess use of electrical energy has a critical effect on environment. Combustion of fossil fuels for example; coal, gas and diesel is considered to be the largest contributing factor to the release of carbon dioxide CO_2 and other harmful gases into the atmosphere. The increasing awareness of environment has encouraged the power utilities to move to the smart grids and renewable energy resources [12]. The enhancement in forecast accuracy helps the operation managers to bring the minimum generating units online in accordance with the peak demand of load [2]. Thus an accurate energy demand prediction can significantly contribute to reduce the unnecessary combustion of fossil fuel and in turns reduces harmful emission of gases to save the environment [12].

The precise and accurate load forecast has a vital impact on power system operations, as economy of operations and control of power systems are quite sensitive to forecast errors [9, 11]. The over and under forecast both would result in increasing the operation cost of power companies. Over forecast results in the form of overcapacity which leads to the waste of money and resources [2]. Under forecast on the other hand can cause power interruptions, which drag down the economy and disrupt the daily activities of the individuals [9]. The enhancement in the accuracy of load demand forecast can save a lot of money and resources [3]. It was observed that a rise of one percent in the forecast error caused an increase of 10 million pounds in operation cost per year for an electric company in the United Kingdom [13].

Besides the importance of load forecast in reducing the generation cost, it is also indispensable for the reliable operation of power systems. STLF can contribute in the estimation of load flows and to make decisions that can prevent overloading. Proper implementations of such decisions improve the network reliability and prevent occurrences of equipment failures and blackouts [2, 7, 14]. It is also useful in the design and implementation of scheduled maintenance plans of power plant equipments, which in turns would enhance the reliability of operation [2]. In a deregulated market, energy demand prediction is also important for contract evaluations and sale and purchase decisions of energy among various power companies [14].

978-1-4673-5074-7/13/$31.00 ©2013 IEEE

The objective of the research is to develop state of the art hybrid ANN-GA models for STLF, and assessment of the performance of these models by applying the actual load data to predict the load of one week in advance. In order to meet this target, the following objectives must be achieved:

- To determine the most appropriate neural network architecture and the values of associated initial weight by implementing GA.
- To develop and implement genetically optimized MLP NN and to train it by implementing back propagation (BP) training algorithm and genetic algorithm.

II. PREVIOUS TECHNIQUES IN STLF

Short Term Load forecast (STLF) is very important in the planning and operation schedules of electric utilities. Several techniques have been investigated to cope up with this problem in the last decades [15, 16]. These techniques can be classified into two categories which are statistical technique and artificial intelligence based technique [3, 17].

A. Statistical Techniques

The statistical techniques normally present the load value by combination of some mathematical equations in which the previous and current values of load and exogenous factors for example the weather and social variables are used [7]. Some of the previous parametric models include regression [5], stochastic time series (autoregressive AR, autoregressive moving average, autoregressive integrated moving average ARIMA) [18, 19], pattern recognition [20] and Kalman filters [17, 19]. The physical transparent nature of these models draws the attention of scientists and researchers and provides a way to understand their behavior. However, they are often criticized for their limited ability to model the nonlinear behavior of load and demand [17]. The other drawbacks associated with these models are the complex layout of the model and heavy computational efforts to produce reasonably accurate results [21].

B. Artificial Intelligence Based Techniques

Modern load forecast techniques based on artificial intelligence (AI), such as expert systems [22], fuzzy logic [23], artificial neural networks (ANN) [1, 6, 11, 21], support vector machines (SVM) [15] and wavelets [24] have been deployed for STLF in the recent past years. Among these techniques ANN has shown very promising results in STLF.

C. Expert Systems (ES)

An expert system (ES) is a computer program, which has the capability to operate as an expert of a certain domain. ES encodes the experience of a domain expert in the form of IF-THEN statements which are called rules. It is capable to interact with the users to provide reasoning and some explanation of the output. Another important feature of expert systems is the continuous upgradation of its knowledge base [21]. It is composed of some facts and rules which are codified in the form of a software. Expert systems are comprised of three main components:

- user interface
- Knowledge base of rules and facts.
- inference engine

An expert system has a unique structure that is different from traditional programs. It is divided into two major parts. First is fixed and independent of the expert system which is known as the inference engine. The second part is variable which is called the knowledge base. To run an expert system, the inference engine executes a reasoning session with the knowledge base similar to human being. The ability of ES of conducting a conversation with users is related to the third component of ES which is called user interface.

From the literature, Rahman et. al., proposed a technique of knowledge base system. In a parameterized knowledge base the load information and inputs are extracted and represented. This technique has been tested on different sites of power utilities in the United States to validate its performance. The load model, rules and the parameters presented in the paper are designed without using any specific knowledge about a particular site.

D. Fuzzy Logic

Fuzzy logic is mathematical in nature and uses Boolean algebraic principles. However its inputs can take the values in the range between 0 and 1 instead of only two outcomes 0 or 1. For instance the temperature may be low, medium and high. Fuzzy inference is the process to produce a mapping from a given input to an output. The mapping provides a foundation from which decisions can be made on the basis of pattern recognition method. After the logical processing of fuzzy inputs, a defuzzification process can be used to produce precise outputs. With such generic conditioning rules and precision in input and output mapping, properly designed fuzzy logic systems can be very robust for STLF.

The advantages of fuzzy logic include the non mathematical model for mapping inputs and outputs and the minimum requirement of precisely processed inputs [7]. The disadvantage of this technique is complex modeling methods for fuzzification and defuzzification processes. It also requires long computing time because of these multiple operations. A variety of fuzzy logic models are reported in the literature for STLF.

E. Multilayer perceptron neural networks

The most popular and frequently reported neural network (NN) in the published work is a multilayer perceptron (MLP) with back propagation (BP) of the error [6, 9].

MLP NN Structure

The general structure of multilayer perceptron neural network (MLPNN) is shown in Fig. 1. First layer is called the input layer, second is a hidden layer and third is an output layer for this multi layer perceptron (MLP). The neurons of the input and hidden layer are connected with the weighted connections called synaptic weights.

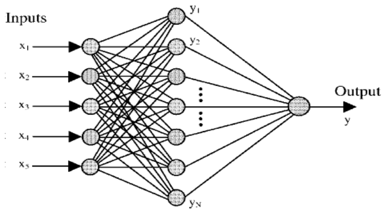

Fig. 1 : MLP NN with three layers.

The strength of an input connection of a neuron is referred as synaptic weight. A neuron model of a single node is shown in Fig. 2. The inputs are linearly summed up after multiplying them by their respective connection weights and then passed through an activation function before passing them on to the next layer. The input output equation of kth neuron in this case can be formulated as:

$$u_k = \sum_{i=1}^{n} x_i w_i \qquad (1)$$

Where u_k is the output of the linear summer, x_i is the i_{th} input and w_i is its respective weight.

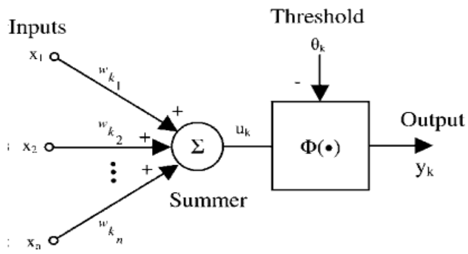

Figure 2 : The neuron model of a single node.

Activation Function

The neural networks pass the output of their layers through an activation function. These activation functions scale the output of the neural network into proper ranges. The most extensively reported activation function in literature is the sigmoid function [4,6].

$$Y_{out} = \frac{1}{1 + e^{-\lambda y_i}} \qquad (2)$$

where y_{out} is the output of the neuron, γ is the sigmoidal gain and y_{in} is the input of the neuron.

There are various forms of sigmoid functions which can be deployed in accordance with nature of problem. Threshold and signum functions are binary in response and work on the principle of McCulloch-pitts model. Other two most commonly used types include, logistic function (ranges the output between 0 and +1) and hyperbolic tangent function (ranges the output between -1 and +1).

ANN Architectures

Many neural network architectures have been developed and reported in the literature such as, feed forward [14], recurrent and functional link with a varying degree of predictive accuracy.

ANN Learning

ANN learning is a process to extract the complex relationship between inputs and output of the network. It is a mechanism of updating the free parameters of the network to improve their performance as the time passes. The general form of learning can be represented by a learning rule:

$$W_{oi}(n+1) = W_{oi}(n) + \Delta W_{oi}(n) \qquad (3)$$

Where $\Delta\, W_{oi}$ is the adjustment of the weight and $W_{oi}(n)$ is the nth time instant.

It helps to calculate the difference between actual and desired output which leads to develop a cost (error) function. In 1986 G.E. Hinton, Rumelhart and R. O. Williams first time proposed Back-Propagation (BP) algorithm which is the most frequently used method for training the neural networks [14]. It deploys gradient descent [4, 6] or conjugate gradient descent algorithms to correct the error function.

$$\Delta W_{oi}(n) = \eta \frac{\partial E(n)}{\partial W_{oi}(n)} \qquad (4)$$

Where E(n) is a suitable cost function, $\partial E(n) / \partial W_{oi}(n)$ is the gradient of E(n) along $W_{oi}(n)$ and ή is a constant called learning rate.

The major drawbacks associated with BP are slow and improper convergence, dependence on initial weight selection which can cause the neural network to get trapped in local minima and the presence of hidden neurons which does not allow a precise picture of the learning process.

F. Evolutionary Algorithms (EA)

An evolutionary algorithm (EA) is a search technique, that uses mechanisms inspired by biological evolution for searching the optimal solution in a complex search space. A number of evolutionary operators such as reproduction, mutation, recombination and selection are implemented in these search strategies. The set of possible solution is first formulated in the form of an initial population. A fitness function then determines the optimal solution that exists in the defined search space (population). Evolution of the search space then takes place after the repeated application of the above mentioned operators. Evolutionary programming techniques such as, genetic algorithms (GA), particle swarm optimization (PSO), artificial immune system (AIS) and ant colony optimization (ACO) have been implemented for optimizing neural networks for STLF application. These techniques produced accurate and fast convergence than BP.

G. Genetic Algorithm (GA)

Genetic algorithm (GA) is a directed random search technique, which was initially discovered in 1970s at University of Michigan by John Holland. The main idea is to design artificial systems retaining the robustness and adoption properties of natural systems. Since the inception, these methodologies were then further improved by other researchers and are now widely used in various fields (business, science and engineering) to solve a variety of optimization problems. GA mimics the biological processes to perform a random search in a defined N-dimensional possible set of solutions. For an optimization problem, it is needed to search and find the best solution in a given search space.

The previous literatures on STLF depicts that genetic algorithms(GA) provide very authenticated approaches specifically for load forecast network structure optimization. As discussed in earlier section, back propagation training based technique can be used to solve various optimization problems. However, as complexity increases, their performance to produce optimal results decays rapidly. On the other hand, genetic algorithm (GA) is regarded as an alternative approach as they are computationally less intensive as compared to back propagation (BP).

GA Characteristics and Operation

The salient features of GA with respect to other function optimization techniques are: they don't need derivative information or other auxiliary knowledge rather they use an objective function, they implement a parallel search in population instead of a single point search, GA implement probabilistic rules not deterministic ones, GA work on encoding of the parameter set rather than parameter themselves, the search towards an optimum solution proceeds not by incremental changes to a single structure but by maintaining a population of solutions from which new structures are created using genetic operators.

GA's work with a set of artificial elements called a population. An individual (string) is referred to as a chromosome, and a single bit in a string is called a gene. GA generates a new population (called offsprings) by applying the genetic operators to the chromosomes in the old population (called parents). Each iterations of genetic operations is referred to as a generation. The evolution process of GA involves following basic steps:

- Initialization of the search node randomly.
- Evaluation of fitness of individuals.
- Selection of two individuals.
- Application of crossover and mutation operators.
- Repetition of the above steps until convergence.

Fitness Function

A fitness function is deployed to evaluate the fitness of an individual in the generation. One of the vital function of GA is to reserve the better schemata, i.e. the patterns of certain genes, so that the off springs may yield superior fitness than their parents. As a result, the cost of fitness function is enhanced through each generation. The fitness function is used to randomly generate and evaluate the strings.

GA Operators

Two types of genetic operators are implemented in genetic algorithms, which are named as crossover and mutation. In the reviewed literature, many forms of the crossover operator are reported such as, two point crossover, multipoint crossover, arithmetic crossover and heuristic crossover. Crossover operator is initialized by selecting the string pairs from the population. Random positions in the string are then chosen and the selected segments are exchanged with the other string that has the similar pattern of partitions. Crossover operators is mainly implemented to explore the new regions in the search space.

The previous literatures on STLF depicts that genetic algorithm (GA) provide very authenticated approaches specifically for load forecast network structure optimization. Back propagation training based technique can be used to solve various optimization problems. However, as complexity increases, their performance to produce optimal results decays rapidly. On the other hand, genetic algorithm (GA) is regarded as an alternative approach as they are computationally less intensive as compared to back propagation (BP). A hybrid approach based on ANN and GA can produce better results in STLF.

III. METHODOLOGY

In order to overcome the deficiencies found in the existing STLF models, a new approach is proposed which integrates the genetic algorithms and artificial neural networks for the development of STLF model. The phases involved in this research work are presented in Fig. 3.

A. Data Preprocessing

The hourly load and weather data for the previous year is collected. As the statistical soundness of the data set plays a vital role in the accuracy of load forecast, so intensive care in the data preprocessing should be carried out.

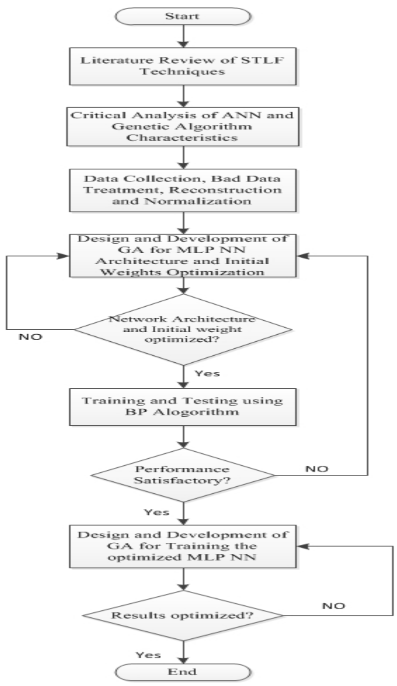

Fig. 3: Phases of proposed research

The first step is to identify the missing or abnormal data and its reconstruction using some effective technique such as fuzzy logic or expert system. The second step involves normalization of the data. The data will be analyzed, interpreted, normalized, and assigned to the model in a simplified manner.

B. Design and Development of Hybrid model

The design, development and implementation of genetic algorithm for the optimization of ANN architecture and initial weight set is an important aspect of this project. This optimized network structure is trained using back propagation technique. The second phase is the complete training of optimized ANN model using genetic algorithm in place of back propagation algorithm. The results of the both techniques are compared. The design and development process is decomposed into the following steps:

- Network architecture optimization using GA
- Initial weight optimization
- Development of GA for overall training of network

C. Network architecture optimization

The most appropriate selection of the number of inputs and hidden layer neurons is the optimization problem in this case. The chromosome length is divided in to two sets. First set represents the number of neurons in the input layer while the second set represents the number of neurons in the hidden layer.

The fitness evaluation function is defined as:

$$fitness = \frac{1}{1+err} \tag{5}$$

Error can be calculated as:

$$err = \sum_{i=1}^{p} \frac{\left|i_f - i_a\right|^2}{p} \tag{6}$$

Where p is the number of samples used during the training process, i_f and i_a are the forecasted output and the actual output and err is the mean square error after a certain number of training epochs.

D. Initial weight optimization

Genetic algorithms is used to provide good initial values for the weights of ANN. Once the exact network architecture is known the number of weighted connection can be determined.

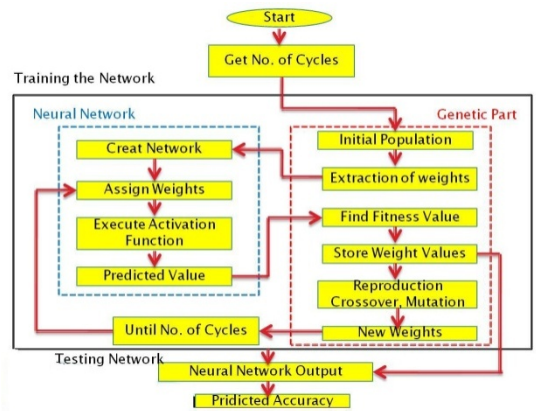

Fig. 4 : Flow diagram for initial weight optimization.

A generic scheme for GA ANN hybrid system is shown in the Fig. 4. Genetic algorithm works for the optimization of ANN initial weight set determination. After the generation of the initial populations, an individual with high fitness (i.e., small output error of the NN) is selected to compete for evolution. The best individual is kept in each generation. The genes of this individual are decoded to give the network weights. Then, the gradient descent technique is employed for the BP algorithm in the NN learning. The complete training of NN with GA is implemented in the next phase.

E. Development of GA for ANN training

The next step is to develop a GA for the complete training of optimized neural network. This algorithm is used in place of BP and employ its own fitness function. The proposed fitness function is based on GA search technique that revises the weights of synaptic connections of NN in a way that leads to minimize the output error.

IV. RESULTS AND DISCUSSION

To overcome the limitations found in previous methods, a hybrid approach that integrates the advantages of ANN and GA is proposed in this research. The genetic algorithm is used for the optimization of ANN architecture and initial weight selection in the first stage. This optimized structure is trained and tested using BP and GA in the second stage. The results of both training methods are compared.

A feedforward neural network based on back propagation training algorithm with gradient descent approach is used in this experiment. A sigmoid activation function is executed in the hidden layer and a linear transfer function is applied in the output layer.

The results obtained from testing the neural network on new data for 24 hours of a day over one week period are presented below in graphical form (Fig. 5). The graph shows a plot of both actual and forecast load in MW against hour of the day for one week. The mean absolute percentage error (MAPE) of 4.35 % is calculated in this technique.

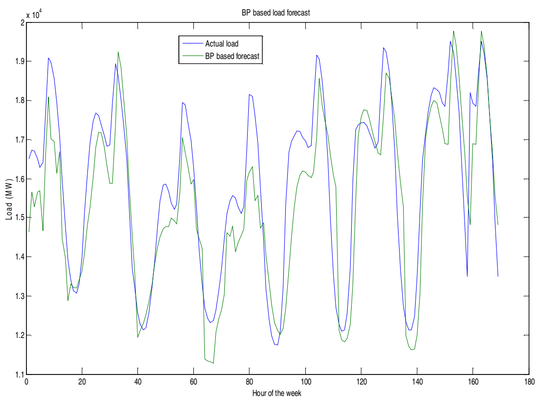

Fig. 5 : BP based forecast result.

In the second experiment GA is implemented to forecast the load for 24 hours of a day over one week period. The graph plot of both actual and forecast load in MW against hour of the day for one week is presented in Fig. 6. The mean absolute percentage error (MAPE) in this method is 1.54 %.

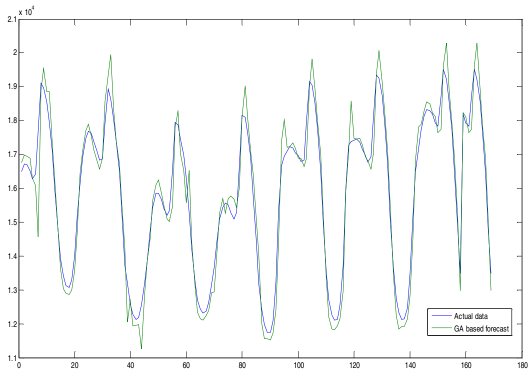

Fig. 6 : GA based forecast results

The overall decrease in the MAPE of almost 3 % reflects that the GA based approach is superior as compared to BP.

V. CONCLUSION

In this paper, the determination of optimized ANN architecture and its initial weights are achieved using genetic algorithm. This optimized network structure is trained with BP in the first stage. Then in the second stage, a new genetic algorithm is proposed for the complete training of optimized neural network in place of BP. The results show that GA based training of ANN produced better accuracy than BP. A considerable reduction in mean absolute percentage error (MAPE) is observed in this approach. The results can be further improved by preprocessing of the input data, increasing the data samples for NN training and increased initial population size for genetic manipulation.

REFERENCES

[1] A. D. Papalexopoulos, S. Hao, and T. M. Peng, "Short-term system load forecasting using an artificial neural network," in Neural Networks to Power Systems, 1993. ANNPS '93., Proceedings of the Second International Forum on Applications of, 1993, pp. 239-244.

[2] F. D. G. George Gross, Senior Members, IEEE, "Short-Term Load Forecasting," Proceedings of the IEEE, vol. 75, December 1987 1987.

[3] H. K. Alfares and M. Nazeeruddin, "Electric load forecasting: literature survey and classification of methods," International Journal of Engineering Sciences & Emerging Technologies, Feb 2012. ISSN: 2231 – 6604 Volume 1, Issue 2, pp: 97-107 ©IJESET.

[4] G. A. Adepoju, "Application of Neural Network to Load Forecasting in Nigerian Electrical Power System," The Practical Journal of Science and Technology, vol. volume 8, 2007.

[5] D. Mrs. J. P. Rothe, A. K. Wadhwani, "Short Term Load Forecasting Using Multi Parameter Regression," International Journal of Coputer Science and Information Security (IJCSIS), vol. vol 6, 2009.

[6] R. W. Nahi Kandil "An efficient approach for short term load forecasting using artificial neural networks," ELSEVIER Electrical Power & Energy Systems, 2006.

[7] D. G. Eugene A. Feinberg, "Book: Title/Chapter: Applied Mathematics For Power Systems, Chapter 12, Load Forecasting".

[8] T. Yalcinoz and U. Eminoglu, "Short term and medium term power distribution load forecasting by neural networks," Energy Conversion and Management, vol. 46, pp. 1393-1405, 2005.

[9] E. F. a. Philippe Lauret a, Rija N. Randrianarivony b, Jean-Franc¸ois Manicom-Ramsamy a, "Bayesian neural network approach to short time load forecasting," Elsevier Energy Conversion and Management, 2008.

[10] T. Senjyu, H. Takara, K. Uezato, and T. Funabashi, "One-Hour-Ahead Load Forecasting Using Neural Networks," Power Engineering Review, IEEE, vol. 22, pp. 70-70, 2002.

[11] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," Power Systems, IEEE Transactions on, vol. 16, pp. 44-55, 2001.

[12] H. Dange and R. Billinton, "Impacts of demand side management on bulk system reliability evaluation considering load forecast uncertainty," in Electrical Power and Energy Conference (EPEC), 2011 IEEE, 2011, pp. 272-277.

[13] Y. Hoseynpoor, T. P. Ashraf, and P. M. Ardabili, "Electrical Load Forecasting Techniques: A Review and Comparison," Journal of Basic & Applied Scientific Research, 2(2)1860-1868,2012, ISSN 2090-4304.

[14] N. S. Pradeepta Kumar Sarangi, Deepak Swain, Dr. R. K. Chauhan, dr. Raghuraj Singh, "Short term load forecasting using neuro genetic hybrid approach: results analysis with different network architectures," Journal of Theoretical and Applied Information Technology © 2005 - 2009 JATIT. .

[15] J. Nagi, K. S. Yap, S. K. Tiong, and M. S. K. Ahmed, IEEE, "Electrical Power Load Forecasting using Hybrid Self-Organizing Maps and Support Vector Machines," The 2nd International Power Engineering and Optimization Conference (PEOCO 2008), Shah Alam, Selangor, MALAYSIA. 4-5 June 2008.

[16] S. Mandal, Urasaki, Funabashi, "A neural network based on several-hour-ahead electric load forecasting using similar days approach," ELSEVIER, Electrical Power & Energy System, 28 (2006) 367 - 373.

[17] H. Chen and A. Singh, "ANN based short term load forecasting in elctricity market," 0-7803-6672-7/01 (C) 2001 IEEE.

[18] M. T. Hagan and S. M. Behr, "The Time Series Approach to Short-Term Load Forecasting," Power Engineering Review, IEEE, vol. PER-7, pp. 56-57, 1987.

[19] I. Moghram and S. Rahman, "Analysis and Evaluation of Five Short-Term Load Forecasting Techniques," Power Engineering Review, IEEE, vol. 9, pp. 42-43, 1989.

[20] A. S. Dehdashti, J. R. Tudor, and M. C. Smith, "Forecasting of Hourly Load By Pattern Recognition A Deterministic Approach," Power Apparatus and Systems, IEEE Transactions on, vol. PAS-101, pp. 3290-3294, 1982.

[21] D. Srinivasan, "Evolving artificial neural networks for short term load forecasting," 0925-2312 (1998 Elsevier Science B.V. All rights reserved. PII: S 0 9 2 5 - 2 3 1 2 (9 8) 0 0 0 7 4 - 5.

[22] H. Ku-Long, H. Yuan-Yih, C. Chuan-Fu, L. Tzong-En, L. Chih-Chien, L. Tsau-Shin, and C. Kung-Keng, "Short term load forecasting of Taiwan power system using a knowledge-based expert system," Power Systems, IEEE Transactions on, vol. 5, pp. 1214-1221, 1990.

[23] A. Jain, M. B. Jain, and E. Srinivas, "A novel hybrid method for short term load forecasting using fuzzy logic and particle swarm optimization," in Power System Technology (POWERCON), 2010 International Conference on, 2010, pp. 1-7.

[24] C. M. Huang "Evolving wavelet-based networks for short-term load forecasting," Generation, Transmission and Distribution, IEE Proceedings", vol. 148, pp. 222-228, 2001.