Asian Journal of Information Technology, 4 (10): 895-900, 2005

Short Term Generation Scheduling Using Cascade Correlation Algorithms with Transmission Constraints

[1]T. Sree Renga Raja, [2]N.S. Marimuthu and [3]R. Shankara Narayanan
[1, 2]Department of Electrical and Electronics Engineering
Government College of Engineering, Tirunelveli, Anna University
[3]TATA Consultancy Services, Chennai, India

Abstract: An artificial neural network based cascade correlation algorithms is investigated for the short term generation scheduling of thermal units considering the real power limit of generators, real power demand, spinning reserve, minimum up and down times of the units. In order to expedite the execution, an Artificial neural network is used to generate a possible unit commitment schedule and a heuristic procedure is employed to modify the unit commitment to achieve a feasible and near optimal solution. The cascade correlation algorithm employs several novel modifications, including the ability to add units when necessary. The results of this method are promising when compared to other existing methods.

Key words: Artificial neural network, cascade correlation, economic dispatch, generation scheduling, optimization and unit commitment

INTRODUCTION

Every modern society mandates on adequate supply of electric energy as economically as possible with a reasonable level of quality and continuity. Short term generation scheduling is currently an area of intensive study, especially with the tendency towards privatization and deregulation of power industry. It is a challenging optimization problem because of the complex interplay amongst many variables.

Most of the operations planning tasks of power system reduce to solving combinational optimization problems. An enormous amount of computation is necessary to solve such problems for large power systems[1]. A neural network is composed of many interconnected multiple single output elements; each constituent is called a neuron. A neuron produces a positive output signal when the sum of its input signals exceeds a certain threshold. Otherwise it yields no output.

Many complex, real world problems are characterized as decision making problems with multiple, conflicting and non-commensurable objectives. The main purpose of the optimal power dispatch problem has so far been mainly confined to minimize the total generation cost of a power system[2]. The multiplicity of objectives and the variety of alternatives, it is often desirable to obtain complete knowledge of the decision maker's global preferred structure explicitly represented by perspective decision model. The advances in ANN make it possible to develop and apply the new methodology and technology to extract decision rules from available information in setting of decision making with multiple objectives.

An ANN consists of many highly interconnected simple and similar processing elements neurons operating in parallel to perform useful computation tasks such as recognizing pre programmed or learned patterns[3]. A crucial property of these networks is their ability to improve the performance by learning new information. Among the various learning algorithms cascade combination as used here.

PROBLEM INFORMATION

The objective of generation scheduling is to minimize the power system operation cost including the cost of fuel for energy generation and starting of process, while satisfying transmission and other system constraints.

Mathematically the optimization problems can be described as follows.

Corresponding Author: T. Sree Renga Raja, Department of Electrical and Electronics Engineering, Government College of Engineering, Tirunelveli, Anna University, Chennai, India TATA Consultancy Services, Chennai, India

$$F=\sum_{j=1}^{T}[\sum_{i\in G}FC_i(P_{ij})+SC_{i,j}] \quad (1)$$

The following are the system and unit constraints, which are taken into account.

- Real power balance Constraint

$$\sum_{i\in G}P_{ij}u_{i,j}=P_{Dj} \quad (2)$$

- Hourly spinning reserve requirements R must be met as

$$\sum_{i\in G}P_{ij}u_{i,j}=P_{Dj} \quad (3)$$

- Real power operating limits of generating units

$$P_i^{min}\le P_{ij}\le P_i^{max}, i\in G, j\in T \quad (4)$$

- Unit minimum up/down (MUT/MDT) time is given as

$$(T_i^{off}, j-1-MDT_i)(u_{i,j}-u_{i,j-1})\ge 0 \quad (5)$$

$$(T_i^{on}, j-1-MUT_i)(u_{i,j-1}-u_{i,j})\ge 0 \quad (6)$$

- The ramp constraints are

$$-P_m^{max}\le P_{mj}\le P_m^{max} \qquad m=1,....,M \quad (7)$$

$$P_{mj}=\sum_{i\in G}k_{mi}P_{ij}U_{ij} \quad (8)$$

- The transmission line constraints are

$$P_{ij}-P_{ij-1}\le UR_i \quad (9)$$

$$P_{ij-1}-P_{ij}\le DR_i \quad (10)$$

ANN IMPLEMENTATION

ANN is a high speed online computational techniques, which are trained through an offline algorithm using example pattern, can provide an output corresponding to a new pattern without any iteration in real time. The cascade correlation network is a constituent algorithm which is used for generation scheduling[4].

An ANN may find it difficult to remember and recognize each pattern. Thus, each generating unit is pre scheduled by a separate ANN by virtue of the fact that this approach is more efficient than training all the units in advance. Each ANN has 24 input neurons corresponding to the 24 hours loads and the output layer has 24 neurons[5]. Load forecast uncertainly imposes and threat to a firm unit commitment decision. Underestimation and overestimation of load can lead to a failure to provide sufficient reserve or lead to an unnecessary large amount of spinning reserve which in turn leads to higher cost.

Cascade correlation algorithm: The cascade correlation approach is a new architecture and supervised learning algorithm for ANNs. Instead of adjusting the weights in a network of fixed topology, cascade correlation begins with a minimal network, then automatically trains and adds new hidden unit one by unit one, creating a multi layer structure[6]. Once a new hidden unit has been added to the network, its input side weights are frozen. This unit then becomes a permanent feature-detector in the network, available for producing output or for creating other more complex feature detector in the network, available for producing outputs or for creating other, more complex feature detectors. The Cascade-Correlation architecture has several advantages over existing algorithms: it learns very quickly, the network determines its own size and topology, it retains the structures it has built even if the training set changes, and it requires no back-propagation of error signals through the connections of the network[7].

The CCA step-by-step

Initial configuration: The algorithm begins with a simple perceptron with N input units and m output units. n and m are chosen on the basis of the problem that the network is to learn[7].

Initial training: The perceptron is trained on the entire training set $\{(Vp,Tp) \mid p = 1, \ldots, P\}$, until the performance of the network is as good as possible. If the desired performance is obtained, the algorithm stops. Otherwise: Start adding hidden units to the network, one by one.

Training of candidates: A pool of candidates for a new hidden unit is generated. This pool emulates a stochastic search in the weight space, which will decrease the risk of inserting a candidate stranded in a local minimum with high error. Each node in the pool of candidates is connected to all input nodes and all previously inserted hidden units. Each of the candidates is trained with the purpose of maximizing some measure of goodness of the candidate.

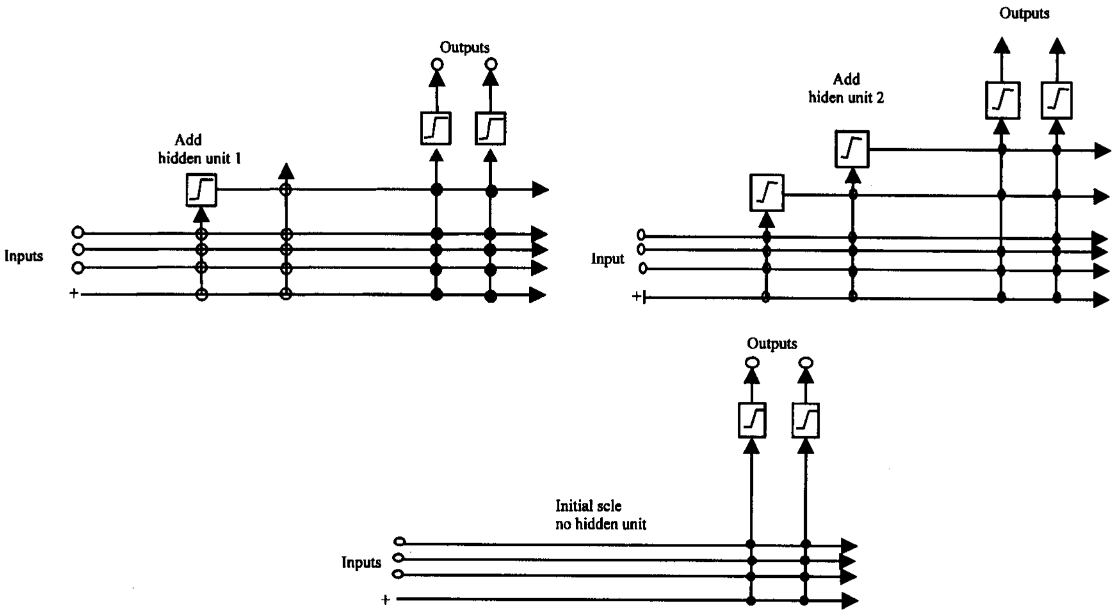

Fig. 1: Cascade architecture initial state and after adding two hidden units. The vertical lines sum all incoming activation. Boxed connections are frozen, X- connections are trained repeatedly

Inserting a new hidden unit: The candidate with the highest score is inserted for real in the network as a new hidden unit. The incoming weights to the new hidden unit are then frozen, i.e. they are not to be changed anymore. The new hidden unit is connected to all output nodes with random weights.

Retraining the network: All the incoming weights to the output units are retrained in order to adjust the weights from the newly inserted hidden unit. If the performance of the network is satisfying after retraining, the algorithm stops. Otherwise: Goto 3.

THEORETICAL FOUNDATIONS FOR CCA

The purpose of inserting a new unit is to reduce the total error of the network. The way the CCA does this is to train the candidate unit so the correlation between the residual error and the output from the candidate is maximized (Fig. 1).

Let X and Y be two stochastic variables. Then the correlation between X and Y is defined as:

$$\mathrm{Corr}[X, Y] = \frac{\mathrm{Cov}[X, Y]}{\sqrt{Var[X]Var[Y]}}$$

It is then possible to show that if X and Y are independent, Corr(X, Y) = 0. On the other hand, if X is proportional with Y, X = cY, where c is a constant, the |Corr[X, Y]| = 1. Thus, if |Corr[V1c, Ei]| = 1 and the activity function is monotonically increasing, it would be possible to find weights so that the error is reduced to about 0. This is easily seen if the activity function of the output unit is the identity function:

$$g(Ui) = Ui.$$

Since the values of Corr[X, Y] lie in the interval [-1,1], the goal of the CCA is therefore to maximise the magnitude of the correlation between the output of the candidate unit to be inserted and the residual error. Empirical results show that performance is improved by using the covariance instead of the correlation. This means that by omitting the normalisation of the covariance, performance is improved. So the interesting term is Cov[X,Y], which is defined as:

$$\begin{aligned}\mathrm{Cov}[X,Y] &= E\left[(X-E[X])(Y-E[Y])\right]\\ &= E\left[(X-\bar{X})(Y-\bar{Y})\right]\\ &= E\left[XY-\bar{X}Y-\bar{Y}X+\bar{X}\bar{Y}\right]\\ &= E[XY]-E\left[\bar{X}Y\right]-E\left[\bar{Y}X\right]+E\left[\bar{X}\bar{Y}\right]\\ &= E[XY]-\bar{X}\bar{Y}-\bar{X}\bar{Y}+\bar{X}\bar{Y}\\ &= E[XY]-\bar{X}\bar{Y}\end{aligned}$$

Both E [X] and X are used to describe the mean of X, the expected value of X and the sample mean value of X respectively.

Table1: Generating Unit Data

Unit	$P_{i\,min}$ MW	$P_{i\,max}$ MW	a_i KS/MW	b_i KS/MW	c_i K$	Bus Number
1	2.40	12	0.025	25.547	24.389	15
2	2.40	12	0.026	25.675	24.411	15
3	2.40	12	0.028	25.803	24.438	15
4	2.40	12	0.028	25.932	27.761	15
5	2.40	12	0.029	26.061	24.888	15
6	4.00	20	0.012	37.551	117.755	1
7	4.00	20	0.013	37.664	118.108	1
8	4.00	20	0.014	37.777	118.457	2
9	4.00	20	0.014	37.890	81.136	2
10	15.20	76	0.009	13.327	81.298	1
11	15.20	76	0.009	13.354	81.264	1
12	15.20	76	0.009	13.381	81.464	2
13	15.20	76	0.009	13.407	81.626	2
14	25.00	100	0.006	18.000	217.895	7
15	25.00	100	0.006	18.100	218.335	7
16	25.00	100	0.006	18.200	218.775	7
17	54.25	155	0.005	10.694	142.734	15
18	54.25	155	0.005	10.715	143.028	16
19	54.25	155	0.005	10.737	143.317	23
20	54.25	155	0.005	10.758	143.597	23
21	68.95	197	0.003	23.000	259.131	13
22	68.95	197	0.003	23.100	259.649	13
23	68.95	197	0.003	23.200	260.176	13
24	140.00	350	0.002	10.682	177.057	23
25	100.00	400	0.002	7.492	310.002	18
26	100.00	400	0.002	7.503	311.910	21

Table 2. Generating Units operating parameters

Units	Min up h	Min Down H	Initial Condition H	UH h	DH H	UR MW/h	DR MW/h	α_i h	β_i h	χ h
1-5	0	0	-1	0	0	48	60	0	0	1
6-9	0	0	-1	1	0	30.5	70	20	20	2
10-13	3	-2	3	2	1	38.5	81	50	50	3
14-16	4	-2	-3	2	2	51	74	70	70	4
17-20	5	-3	5	3	2	55	78	150	150	6
21-23	5	-4	-4	4	2	55	99	200	200	8
24	8	-5	10	5	3	70	120	300	200	8
25-26	8	-5	10	8	4	50.5	100	500	500	[illegible]

$SC_{i,j} - U_{i,j}\{1-U_{i,j-1}\}[\alpha_i + \beta_j[1-\exp\{T^{off}_{i,j}/\tau_j\}]]$

Table 3: Hourly load demand and bus load as percentage of total load

Hour	Load,MW	Hour	Load, MW	Hour	Load,MW	Bus number	Bus load, %	Bus number	Bus load, %
1	2000	10	2257	19	2539	1	3.8	10	6.8
2	1847	11	2308.5	10	2488	2	3.4	13	9.3
3	1744	12	2334	21	2411	3	6.3	14	6.8
4	1693	13	2308.5	22	2360	4	2.6	15	11.1
5	1642	14	2257	23	2231.5	5	2.5	16	3.5
6	1643	15	2231.5	24	2077.5	6	4.8	18	11.7
7	1693	16	2231.5	-	-	7	4.4	19	6.4
8	1795.5	17	2334	-	-	8	6.0	20	4.5
9	2052	18	2565	-	-	9	6.1	-	-

Table 4: Results of ANN with the Back Propagation and Cascade Correlation

Sample No	ANN- BP $	CPU, S	ANN-CC $	CPU, S
I	441128	98	438702	96
II	439905	96	433821	90
III	438091	98	432362	93
IV	441008	96	425724	92
V	438917	94	434240	91
VI	439597	99	434545	89
VII	439072	94	434344	90
VIII	435138	98	429855	95

The minimization cost function is transformed into a maximization problem by the reciprocal of the cost function.

$$F_{GA} = \sum_{j=1} F(j) + \alpha_1 SC(j) + \alpha_2 SDC(j) + \alpha_3 FTR(j) + \alpha_4 FTD(j) + \alpha_5 VTL(j) \quad (11)$$

where $\alpha_1, \alpha_2, \ldots, \alpha_5$ are penalty factors, SC(j), SDC(j), FTR(j), FTD(j) and VTL(j) are the penalty terms, which are utilized either when a unit up does not comply with the minimum up/minimum down time, or fails to meet the reserve and the load demand.

Table 5. Results of unit scheduling generated by CCA with Transmission Line Constraints

Units	Hours (1-24)																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	0
2	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	1	1	0	1	0	0
3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	1	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0
5	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0	1	0	0
6	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
16	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0
24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

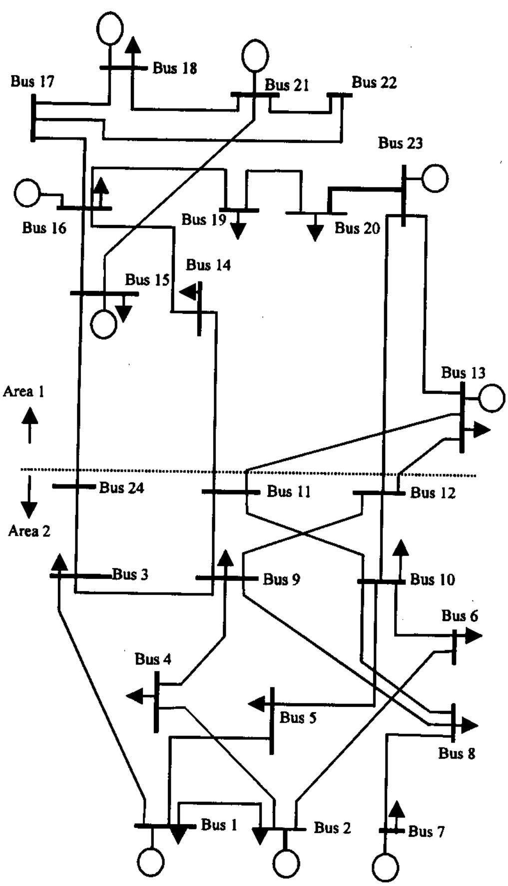

Fig. 2. IEEE 24 Bus System

DISCUSSION

The proposed techniques are applied to a medium size system consisting of 26 units. Its transmission network consists of a 24 bus location connected by 38 lines as shown in Fig. 2. To facilitate comparing the results of the proposed techniques to other recently reported approaches utilizing Back Propagation (BP)[11] in this case, the transmission line and ramp rate constraints are taken into consideration. The generating unit data, generating operating data and hourly load demand data are obtained from a standard IEEE 24 bus system as shown in the Tables 1 to 3. In this case, the merit order method is used for dispatching the load amongst the committed units, due to the fact that this method gives better results and uses less computational time as shown in Table 4.

CONCLUSIONS

The cascade correlation algorithm based on ANN is proposed to effectively solve the problem of short-term generation scheduling. The proposed techniques are tested on the IEEE reliability test system consisting of 26 generating units, taking into consideration the system and unit constraints. Two problem formulations are introduced according to the type of fuel cost and start up functions and the group of constraints. With regard to ANN, a long time can be expanded on offline training of the network as ANN accumulates knowledge during offline training from the given input/ output data pairs. However once the network is completely trained, the online response would be very fast compared to analytical programming techniques. It is also expected that the proposed technique can be applied to different problem dimension and could score more favorable compared with analytical techniques. As a next step, the proposed technique can be extended and environmental constraints also taken as one of the constraints.

REFERENCES

1. Wang, S.J., S.M., Shahide hpour, D.S. Kirschen, S. Mokhtari and G.D. Irisarri, 1995. Short-term generation scheduling with transmission and environmental constraints using an augmented Lagrangian relaxation. IEEE Trans. Power sys., 10: 1294-1301.
2. Wood, A.J. and B.F. Woolenberg, 1996. Power Generation Operation and Control Jhon.Wiley and Sons, New York.
3. Sasaki, H., M. Watanabe. and R. Yokoyama, 1992. A Solution method of Unit Commitment by Artificial Neural Network. IEEE Trans. On Power System , 7: 974-981.
4. Fahlman, S.E. and C. Lebiere, 1990. The Cascade-Correlation Learning Architecture in Neural Information Processing Systems 2, D. Touretzky, (Ed.) Morgan Kaufmann Publishers, Colorado, pp: 524-532.
5. Kulkarni, P.S., A.G., Kothari and D.P. Kothari, 2000. Combined economic and emission dispatch using improved back propagation neutral network, Electr. Power Components Sys., 11: 31-44.
6. El Desouky, A.A. and M. Elkateb, 2000. Hybrid adaptive techniques for electric-load forecasting using ANN and ARIMA. Proc. Gene.Trans. Dist., 147: 213-217 .
7. Rivest, F. and T.R. Shultz, 2002. Application of Knowledge-based cascade-correlation to vowel recognition. In proceedings of the 2002 International Joint Conference on Neural Networks, pp: 53-58.