DOI: http://dx.doi.org/10.26483/ijarcs.v9i2.5679

ISSN No. 0976-5697

Volume 9, No. 2, March-April 2018

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

APRIORI BASED MACHINE LEARNING IN POWER DISTRIBUTION NETWORK

Jasreen Kaur
Computer engineering and Technology Department
Lovely Professional University
Phagwara, India

Abstract: Power Distribution Networks (PDNs) are responsible for delivery of electricity from transmission substations to individual belonging. PDNs are still an open area of research and demand efficient load balancing algorithm. Although initially loads are balanced among three phase network but drift into unbalance state meeting ever growing and fluctuating demands. This imbalance in loads at different phases of system affects tool utilization, voltage ranges system stability and security .The comprehensive review has shown that the use of machine learning in not considered in PDNs. The issue of scalability has been ignored. This paper proposed Apriori based machine learning algorithm to predict the loads and schedule it in the optimum way. The comparisons have clearly shown that the effectiveness of the proposed technique over the existing one.

Keywords: load balancing; machine learning; Power Distribution Networks; phase swapping; feeder reconfiguration.

I. INTRODUCTION

Electric Power seems to have become the part of every household via massive distribution systems through which electric power moves significant distance from origin to each customer through service mains. Often there is uncertainness of loads among feeder w.r.t time. There is always a need to balance the load among the feeder to prevent this imbalance as it vulnerable to system security, equipment utilization and voltage drops. Phase imbalance increases power losses due to quadratic nature of losses [2]. It also improves line safety measures pertaining to labours and increases overhead on power program. At the same time many loads are vulnerable to imbalance and cannot further precede proper procedure. Fig.1 describes the basic architecture of the PDNs. The supply of electric power is done through distributed substation. Substations are connected to feeders which in turn are connected to transformer. Distribution transformer act as terminal voltage conversion dropping down voltage to level utilized by customer.

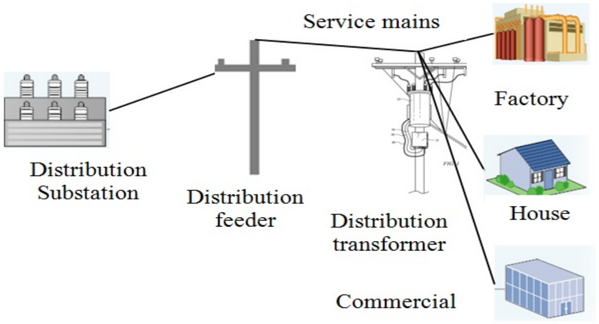

Figure 1. Architecture of PDN

Load balancing in PDNs is an act of circulating the loads among different feeders such that no feeder remains overwhelmed. Even if loads are initially balanced among three phase distribution system yet drift into unbalance state meeting ever-growing and fluctuating demands. The imbalance in load at different phases of system affects equipment usage, voltage levels and system security [3]. Phase imbalance increases power losses due to quadratic nature of losses [2]. It also improves line safety measures pertaining to labours and increases overhead on power program. At the same time many loads are vulnerable to imbalance and cannot further precede proper procedure.

Prediction of load prior the demand is an important and challenging problem. This can only be done through machine learning. Machine learning is the art by which computer learn and act without being explicitly programmed. Here the learner is a system which predicts the loads for each feeder and then checks the difference in predicted data set an actual one to

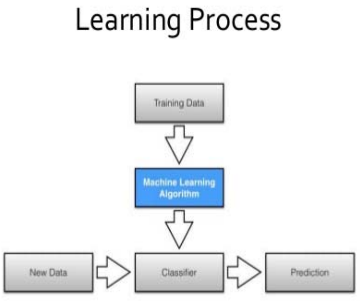

Figure 2. Machine Learning Process

II. LITRATURE SURVEY

In [4] simulated annealing is used for large scale non linear integer encoding to find optimum option for phase balancing with high computation time. Using expert system in [5] and immune algorithm in [6] rephasing is resolved by removing expense regarding labor as well as customer service is reduced. In [7] a combination of fuzzy logic and network Raphson used to make phases balanced by on/off different switches subject to constraint that for each load minimum one switch to be closed. In order to balance the load among the three phase system of PDNs DILB method can be used which enhance the capabilities of the network by the use of new digital power meters placed in installation panels that are further connected to local computer network via Internet. At each node an intelligent rotary cam switch is added and microcontroller to drive that switch. every microcontroller consistently determine the current and voltage in its node at predefined sampling rate. Subsequently most of the nodes should synchronize with each other and shift the rotatary cam switch in locations which will result in minimum imbalance at root node. Results show that this method provides minimum load disbalance and power losses [1].In different areas data mining is used to solve the load balancing problem such as distributed heterogeneous environment[9] cloud computing[10] grid computing [11].Apriori algorithm is very much effective in finding frequent use item set and association rules .In[12] data mining is used to enhance the performance in multi core architecture using Apriori algorithm.

III. Proposed Methodlolgy

Apriori Machine Learning is a an efficient association rule based learning technique based on finding an association rules between frequent used item set in the given dataset. Modeling such a technique for power distribution network is an intriguing region of exploration and great deal of concepts have been brought ahead to intimate this goal. Machine Learning is such a well-known area, which has been freshly believed to deal with complex problems. Figure 5.1 depicts the outline of an algorithm.

In the proposed algorithm we start with first initializing the data center requirement. In this stage the servers are assigned the requirements using normal distribution. Apriori algorithm is then applied to the given dataset to find the best communication matrix from their past experience or reading to find the best arrangement in which loads are assigned to servers such that the servers remain balanced. After this we develop a communication cost matrix which decides how much cost is engaged in transferring load from one to another. Genetic algorithm follows this step to find the optimal solution of the problem in the minimal time as it uses probabilistic rules over deterministic one. It follows the genetic algorithm steps then select the solution if it balances the feeder/ server such that in three phase system the sum of all phases sum up to zero. After that we evaluate parameter such as overall gain and total disbalance in the current system.

For evaluation of the new proposed time three parameters are considered i.e. execution time, overall disbalance and overhead cost. Each parameter is further evaluated when we change number of servers and power distribution value respectively. Evaluation is further categorized into two parts when we have taken all the resources of same capacity i.e for homogenous resources and another one for heterogeneous resources. Three metrics for evaluation are explained below:

1.Execution Time

It is a time taken by system executing the task here task is swapping the loads over a three phase electric system. It is summation of communication time and computation time. If loads over network is balanced, execution time will be less and more useful the system will be.

$$Execution\ Time = Communication\ Time + ComputationTime$$

2.Load Balancing Gain

This metric can be defined as how the work load is divided among the phases in other words variation of load among different phases.

$$G = \frac{\sum_{i=1}^{n} U_i}{n}s.$$

Where i=1,2,3...........n Servers and Ui=Utilization

3.Overhead Time

It is defined as extra time required by the system to attain the required goal. It is also called communication time needed to make a coordination link with the different phases of feeders to which power is to be transferred in order to balance it. Our main goal is to decrease the overhead time.

$$D = \frac{\sum_{j=1}^{n} \sum_{i=1}^{n} U_i - U_j}{n}$$

A. Algorithm

Step1: Deploy Power Distribution Network and load the data
Step2: Assign Server Requirements using normal distribution.
Step3: Apply Apriori Algorithm
Step4: Deploy Communication cost matrix
Step 5 Apply Genetic algorithm
Step 6: Apply crossover
Step 7: Apply mutation
Step 8: Selection
Step 9: if Stopping criteria fulfilled then go to step 10 else step 5
Step 10: return Final results

B. Flowchart

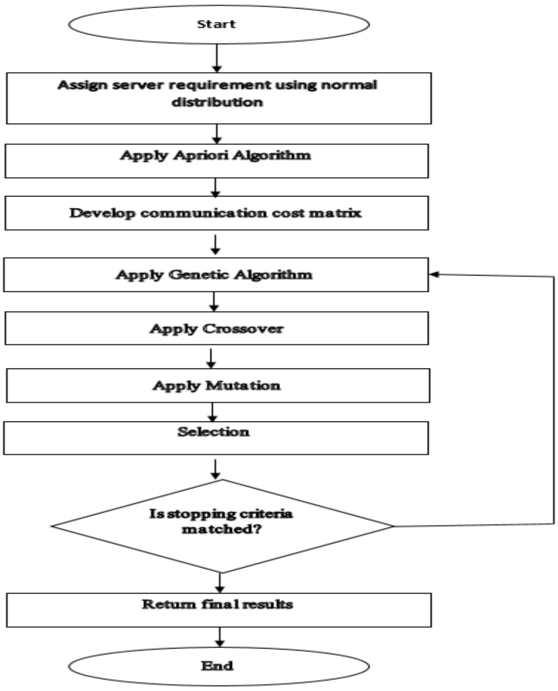
Start
Assign server requirement using normal distribution
Apply Apriori Algorithm
Develop communication cost matrix
Apply Genetic Algorithm
Apply Crossover
Apply Mutation
Selection
Is stopping criteria matched?
Return final results
End

Figure 3. Flowchart of proposed algorithm

IV. Experimental Setup

The experiments are done by considering the Microsoft Windows 8.1 with Intel i5 processor along with 2 GB NVidia graphics card and 8 GB RAM. To test the validity of proposed method, it is implemented using MATLAB of version 7.10 using signal processing toolbox. MATLAB is a high level and interactive environment for analysis of data, numerical computations, and visualization of data and development of algorithm. Table 1 is representing various symbols used in the simulation work along with their respective range. However our work is not limited to these stated values, these values are only taken for better evaluation purpose

Table I. Notations

Symbol	Meaning	Value
ns	No.Of Servers	60-105
pd	Power Distribution	40-85
$KVMs$	Kilowatt per Machines	10
P	Present initial storage space	5Gb storage space,5Mbps bandwidth,5MIPS user cycles
uD	User Demand	0-5MIPS
BD	Bandwidth Demand	0-5Mbps
SD	Storage space Demand	0-5Gb
CT	Communication Traffic	0-x based on topology
g	Load Balancing Gain	0-∞
D	Overall Disbalance	0-∞
UBS	uD, BD, SD	User demand ,Bandwith Demand,storage Space
F	Flag	0-1
OPVM	Old PVM Flag	0-1
TS	Temprory server	0-1
CI	Cross Individual	0-1
PN	Population Number	30

.

V. Experimental Results And Performance Evaluation

This section provides cross validation and verification of existing and proposed techniques with respect to some well defined performance evaluation criteria The proposed method is evaluated over different performance metrics such as execution time, load balancing gain, overhead time ,overall disbalance with respect to number of servers and power distribution in comparison to existing technique.

A. *Performance evaluation based on Execution Time*

Execution time changes with change in number of servers. Here we have taken the value of number of servers from 60 -105.

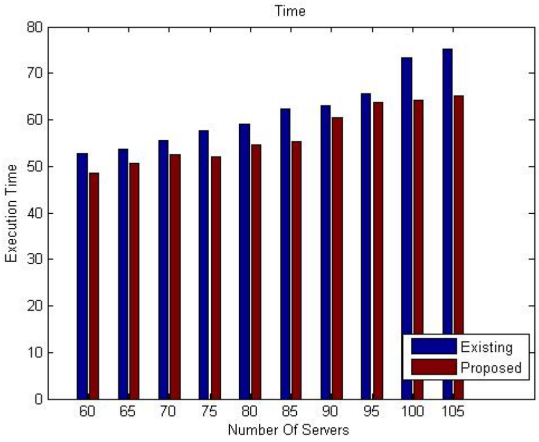

Figure 4. Comparision of execution Time Of existing and proposed over number of servers in homogenous environment

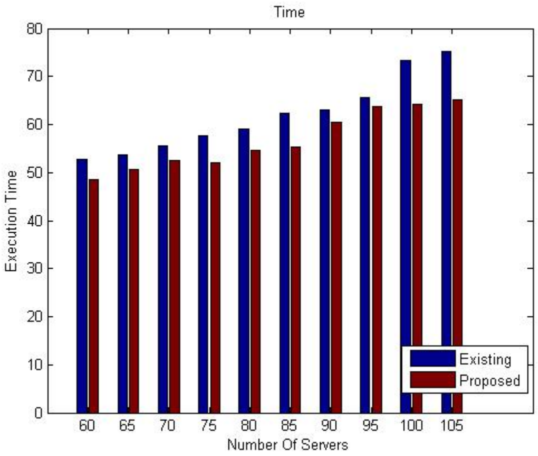

Figure 5. Comparision of execution Time Of existing and proposed over number of servers in homogenous environment

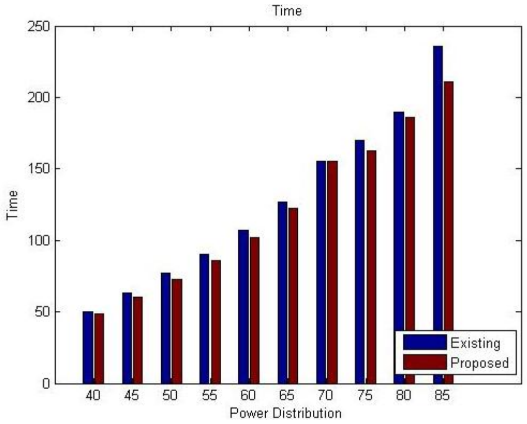

Figure 6. Comparision of execution Time Of existing and proposed over Power Distribution in homogeneous environment

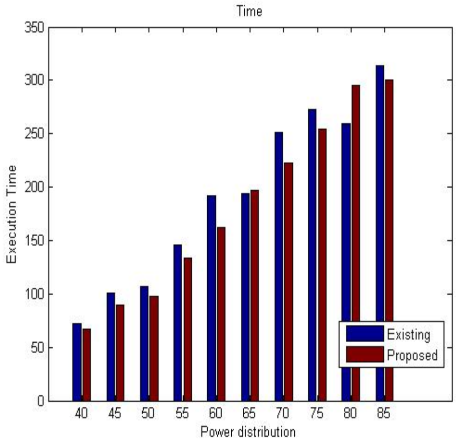

Figure 7. Comparision of execution Time Of existing and proposed over Power Distribution in hetrogenous environment

Figure. 4 and Figure 5 shows the comparison among Existing Dynamic Intelligent Load Balancing (DILB) and Proposed Machine Learning Based DILB over execution time w. r. t No of Servers for homogenous and heterogeneous environment. Figure. 6 and Figure 7 shows the comparison among Existing Dynamic Intelligent Load Balancing (DILB) and Proposed

Machine Learning Based DILB over execution time w. r. t Power Distribution for homogenous and heterogeneous environment. Bar graph clearly shows that the time evolved to swap the load over different phases has decreased in case of the proposed as compared to the Existing DILB. It is confirmed that the proposed algorithm is comparatively better than the existing techniques. This decrease represents improvement in the speed of swapping or reassignment of loads.

B. *Load Balancing Gain*

This metric can be defined as how the work load is divided among the phases in other words variation of load among different phases.

$$G = \frac{\sum_{i=1}^{n} U_i}{n} s.$$

Where i=1,2,3...........n Servers and Ui=Utilization

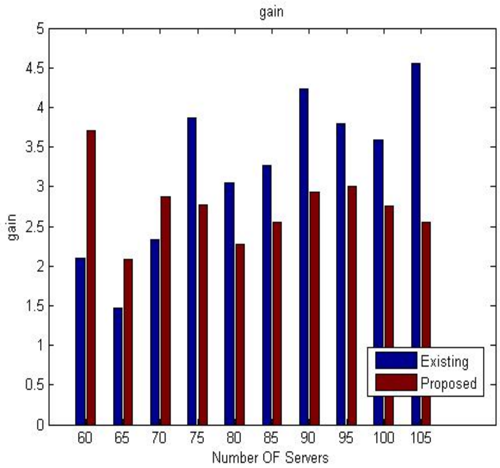

Figure 8. Comparision of Overall Gain Of existing and proposed over number of servers in homogenous environment

.

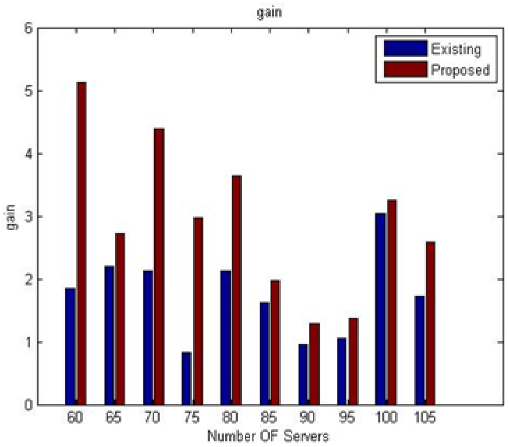

Figure 9. Comparision of Overall Gain Of existing and proposed over number of server in heterogenous environment

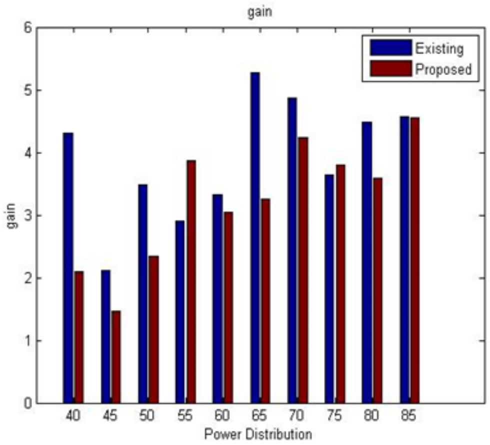

Figure 10. Comparision of Overall Gain Of existing and proposed over over Power Distribution in homogenous environment

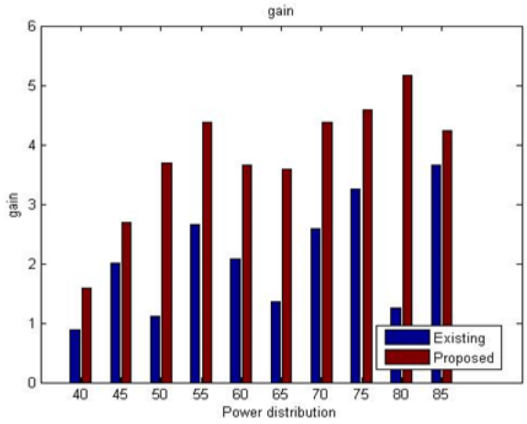

Figure 11. Comparision of Overall Gain Of existing and proposed over over Power Distribution in hetrogenous environment

Figure 8 and Figure 9. shows the comparison among Existing Dynamic Intelligent Load Balancing (DILB) and Proposed Data Mining Based DILB over gain w. r. t. No of Servers for homogenous and heterogeneous environment. Figure. 10and Figure.11 shows the comparison among Existing Dynamic Intelligent Load Balancing (DILB) and Proposed Data Mining Based DILB over gain w. r. t. Power Distribution for homogenous and heterogeneous environment Bar graph clearly shows that the increase in gain when we use machine learning to swap the load over different phases in case of the proposed as compared to the Existing DILB. It is confirmed that the proposed algorithm is comparatively better than the existing techniques.

C. *Overhead Time*

It is defined as extra time required by the system to attain the required goal. It is also called communication time needed to make a coordination link with the different phases of feeders to which power is to be transferred in order to balance it. Our main goal is to decrease the overhead time.

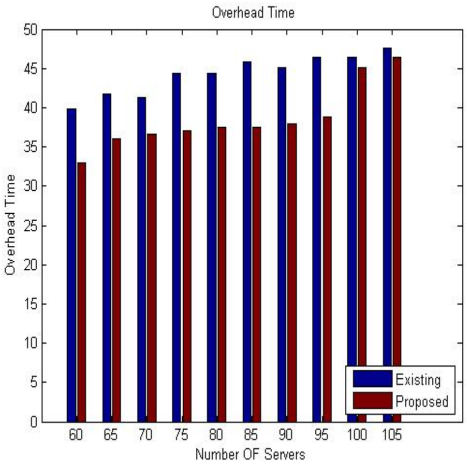

Figure 12. Comparision of Overhead Of existing and proposed over number of servers in homogenous environment

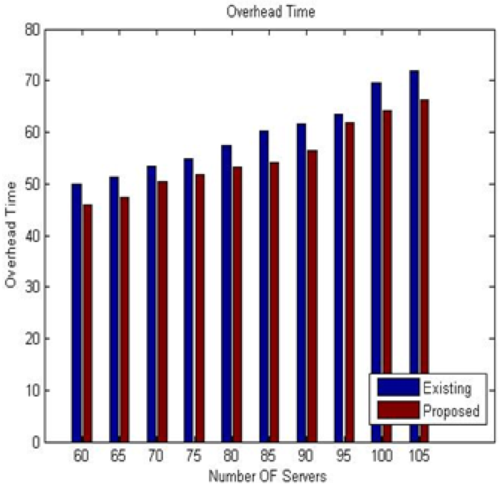

Figure 13. Comparision of Overhead Of existing and proposed over number of servers in hetrogenous environment

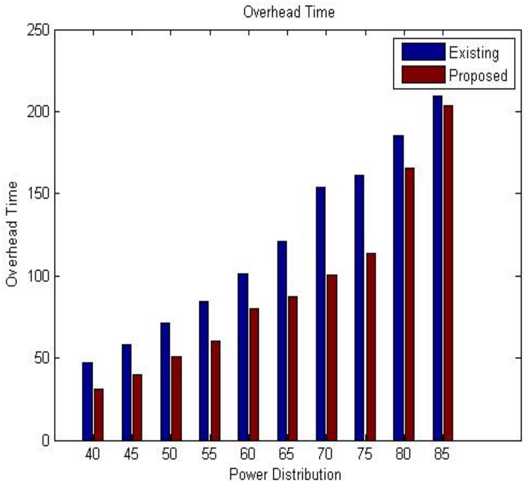

Figure 14. Comparision of Overhead Of existing and proposed over Power Distribution in homogenous environment

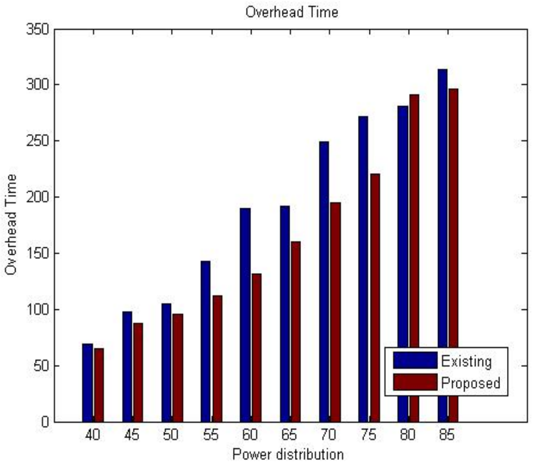

Figure 15. Comparision of Overhead Of existing and proposed over Power distribution network in hetrogenous environment

Figure12., Figure 13 shows the comparison among Existing Dynamic Intelligent Load Balancing(DILB) and Proposed Data Mining Based DILB with respect to overhead in homogenous and heterogeneous environment respectively. Bar graph clearly shows that the time evolved to swap the load over different phases has decreased in case of the proposed as compared to the Existing DILB. It is confirmed that the proposed algorithm is comparatively better than the existing techniques. This decrease represents improvement in the speed of swapping or reassignment of loads.

VI. Conclusion

After reviewing the literature it was found that there is no technique which considered machine learning or data mining to enhance the performance of PDN. The scheduling of loads over different phases at the time of overload is still an important issue in concerned field. Reconfiguration the whole network is regarded as expensive and time consuming job in which system has to be shut down. Phase swapping outlooks these drawbacks. But still it does not incorporated the idea machine learning for load prediction and balancing. We thereby presented a novel approach in solving load balancing problem by the use of Apriori machine learning algorithm. So the values we obtained from simulation environment clearly shows the improvement in performance of PDNs over existing technique. Moreover reduction in overhead and execution time clearly shows the superiority of proposed algorithm over existing literature..

VII. References

[1] Vladimir Zdraveski, Mirko Todorovski, Ljupco Kocarev, "Dynamic intelligent load balancing in power distribution networks", in International Journal of Electrical Power & Energy Systems on , vol. 73, pp. 157-162, May 2015 ,ISSN0142-06158

[2] Kuo C-C, Chao Y-T,"Energy managment based on aim/fm/gis for phase balancing application on distribution systems", in EnergyConcers Manage on ,vol.51,No.3, 2010

[3] Gupta, N.; Swarnkar, A.; Niazi, K.R., "A novel strategy for phase balancing in threephase four-wire distribution systems," in Power and Energy Society General Meeting, 2011 IEEE , vol., no., pp.1-7, 24-29 July 2011

[4] Jinxiang Zhu; Bilbro, G.; Mo-Yuen Chow, "Phase balancing using simulated annealing," in Power Systems, IEEE Transactions on , vol.14, no.4, pp.1508-1513, Nov 1999

[5] Chia-Hung Lin; Chao-Shun Chen; Hui-Jen Chuang; Ming-Yang Huang; Chia-Wen Huang, "An Expert System for Three-Phase Balancing of Distribution Feeders," in Power Systems, IEEE Transactions on , vol.23, no.3, pp.1488-1496, Aug. 2008

[6] Huang, M.-Y.; Chen, C.-S.; Lin, C.-H.; Kang, M.S.; Chuang, H.-J.; Huang, C.-W.,"Three-phase balancing of distribution feeders using immune algorithm," in Generation, Transmission & Distribution, IET , vol.2, no.3, pp.383-392, May 2008

[7] W.Siti, A. Jimoh and D. Nicolae,"Phase load balancing in secondary distribution network using fuzzy logic and combinatorial; optimization based on newton raphson," in Proc. Intelligent Data Engineering and Automated LearningIDEAL, vol 5788, pp. 92-100, 2009.

[8] S. Senhadji, S. Khiat and H. Belbachir, "A dynamic load balancing strategy for generating association rule mining under a grid environment," Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), 2012 6th International Conference on, Sousse, 2012, pp. 976-981

[9] Tlili Raja and Slimani Yahya ," A Hierarchical Dynamic Load Balancing Strategy for Distributed Data Mining,"in International Journal of Advanced Science and Technology Vol. 39, February, 2012

[10] Priyanka R," Load Balancing In Public Cloud Combining The Concepts Of Data Mining And Networking", International Journal of Research in Engineering and Technology, Vol.3 No.3,3 May 2014,pp167-171.

[11] Saeed Sharifian, Seyed A. Motamedi, and Mohammad K. Akbari.," A predictive and probabilistic load-balancing algorithm for cluster-based web servers". Appl. Soft Comput. Vol.11,No.1, January 2011,pp. 970-981

[12] Kun-Ming Yu, Sheng-Hui Liu, Li-Wei Zhou, and Shu-Hao Wu., "Apriori-based HighEfficiency Load Balancing Parallel Data Mining Algorithms on Multi-core Architectures".

[13] M.Eugin Lilly Mary, Dr.V.Saravanan, “Predictive Load Balancing For Data Mining In Distributed Systems ”, Journal of Theoretical and Applied Information Technology, Vol. 53 No.1, , 10th July 2013