International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org
Volume 3, Issue 5, September-October 2014 **ISSN 2278-6856**

Using Hybrid Model of Turing Band Method and Feed Forward Neural NetworkImprovingPerformance Accuracyof Distributed Web System

Miss. Jyoti J. Dhayagude[1], Dr. Pradeep K. Deshmukh[2]

[1]Student, Dept. of CSE, Pune, India

[2]Dr, Dept. of CSE, Pune, India

Abstarct

Recently the prediction method was introduced for forecasting the web performance by using the innovative application of Turing Band (TB) geostatistical experiment. However this method is having major limitation which is required to be rectified further. The forecasting accuracy of this proposed method is poor. In addition to this, this method was only assuming the input data in linear form to forecast the web performance, however by considering the real time environment, the input data may be vary and non-linear as well. Therefore this approach fails for different measurement data as well as prediction horizon lengths. The later limitation is out of scope of this paper, we are presenting the extended TB method to improve the accuracy of forecasting. We have studied many methods those are used to improve the forecasting accuracies. The artificial neural network (ANN) is one of the best methods to improve the forecasting accuracy. In this paper we are presenting the hybrid model of TB and Feed Forward Neural Network (FFNN) together for improving the overall accuracy performance. The experimental study is performed using the real time web log dataset and comparative analysis of forecasting accuracy is done in between TB and Hybrid TB method.

KEYWORDS:- Forecasting accuracy, feed forward neural network, geostatistics, spatio-temporal prediction, Web of Things (WoT), Web performance forecasting.

1. INTRODUCTION

Now days it is observed that there is tremendous growth of using distributed computer systems (DCSs) and its development word wide. The main development area of DCSs are manufacturing and industrial applications over the world wide web (WWW) architecture such as service oriented architecture as well as web related QoS aware systems. This is very helpful to forecast the performance web system [1].

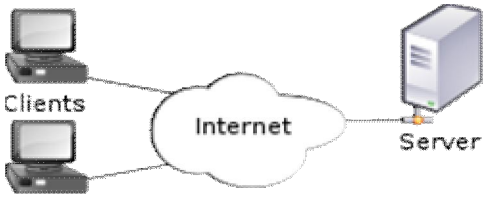

Figure 1: System Architecture.

Numerous amounts of electric power quality assessment and index to the same point of common coupling connected requires simultaneous measurements in all lines is growing evidence in the literature, the measuring system based on digital signal processing techniques during recent years anddigital systems interconnecting and data exchange capability and more attractive and cost effective power quality applications these systems are made that performance Growth.Moreover, the availability of a world-wide, low-cost, and public-domain interconnection system, the Internet, is pushing the evolution of the remote measurement systems, where the measurement results provided by in-field measurement systems are collected and stored by a central unit, toward the distributed measurement systems, where different systems, located in different places, share the same data in order to perform a measurement. It is known that the major drawback of these systems is the lack of synchronization of the shared data, due to the variable and unpredictable throughput of the net, which may affect the uncertainty. There are many methods each method suffered from limitations, but Web systems, to predict the performance of the prediction are presented. In fact there wrong Web performance problems are two possible solutions to overcome the first solution focuses on improved communication protocol and another solution for efficient distributed computer systems Web based performance prediction methods. In [1], authors introduced the new TB geostastical based web system forecasting performance system. However as we stated in abstract of this paper, the existing method presented in [1] is suffered from limitations like poor forecasting accuracy and cannot process all kinds of measurement data and other parameters. It ranges in this paper to new hybrid method combining two concepts based TB [1] and [2], aimed at improving the accuracy of forecasts FFNN's concepts are introduced in a variety of ways the next section II. Web system presented on prophecy and ANN are more than literature survey. In section III the proposed approach and its system block diagram is shown. In section IV we are presenting the current state of implementation and results achieved. Finally conclusion and future work is predicted in section V.

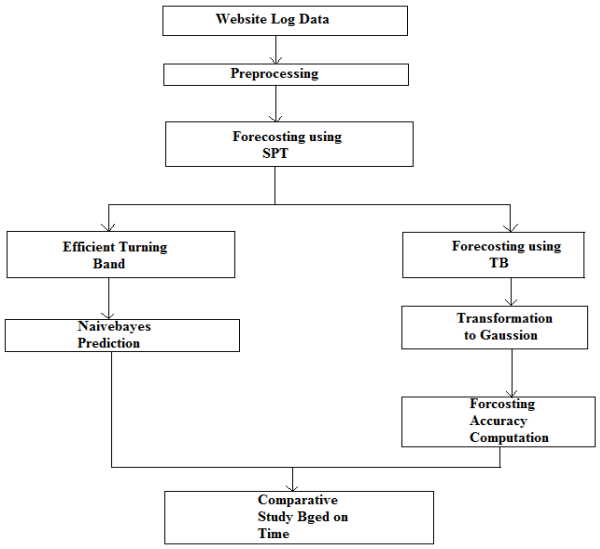

Figure 2: Block diagram of system

2. LITERATURE SURVEY

In this section we are presenting the different methods those are presented to solve inaccurate performance prediction methods.

- In [3], M. Ulieru and the latest trends in today's complex interplay of dynamic intelligent technologies and services presented and supported by a network of interdependent ecosystem of the world converging Grobbelaar revealed. our efforts in these prototype and upcoming future events to meet the challenges of tomorrow in preparation for industrial ecosystems for use with a solid foundation in setting are detailed.
- In [4], F. Tao, Hu and Zhao, y z Zh d. formulation presents author minimize implementation time and cost, and reliability to maximize a MO-MRSCOS are presented for an issue. basic resource service mixed mode (RSCM) are described for CRS, and principles in a complex sequence of a simple translation of the RSCM presents to solve process and MO-to simplify the complexity of MRSCOS. Particle Swarm (PSO), optimization based on the principles of a new structure and optimum selection method, MGrid resources proposed again.
- The PSO follows a collaborative population-based search, which models based on the social behavior of bird flocking and fish schooling.
- In [5], T. Cucinotta, Mancina, g. f. Anastasi, g. Lipari, l Mangeruca, R. Checcozzo, and f. Rusina writers conformance testing service-oriented testing framework for interconnection technology. Structures established on previous research and market analysis-based research track. Conformity test is a case study, which tested the network on Terminal architecture (NoTA) NoTA focuses on concentrating on running on terminal equipment application framework that communication between Facilitate a service oriented interconnection architecture.
- In [6], L. Borzemski, k Zatwarnicka an Adaptive decision Zatwarnicki and a..Framework, namely a fuzzy logic and neural network resource request within a geographically distributed fully replicated Web sites for Adaptive and intelligent dispatching algorithm to deploy two machine learning technology applications. Our approach and algorithm can be used in content distribution and delivery networks where due to content replication each Web server can respond to the client request. Intelligent redirection mechanism redirects the request from the originating site's server to the “best” content source to deliver the content in the fastest manner.
- In [7], Q. Wang and s. Gopalakrishnan introduced real-time industrial control systems scale, single-real time local area network (LAN) is no longer enough; Instead, we have several real-time LANs real time wide area network (WANs) to merge into real-time switch. However, nowadays commercially-off-the-shelf wan switches rather than best effort Internet traffic real-time traffic are designed for. To address this problem, we propose a real-time crossbar switch design that minimally modifies, and even simplifies the de facto industrial standard switch design of iSLIP. Specifically, we change the iSLIP request-grant-accept negotiation to deterministic grant. The switch runs periodically with an M cell-time clock-period.
- In [8], Onat, T. Naskali, E. Parlakay, and O. Mutluer authors described MBPNCS and then introduce a stability criterion. This is followed by computer simulations and experiments involving the speed control of a dc motor. The results show that considerable improvement over performance is achieved with respect to an event-based NCS.
- In [9], Cristaldi, a. l. Ferrero, c. Muscas, s. Salicone and r. Tinarelli delivering electric power quality measurements, the authors measurement systems introduction and how to use an appropriate technique, average an Internet connection through the data transmission can be reduced to potentially harmful effects thus distributed among various units of measurement systems and avoid costly sync at the end. Measurement uncertainty of delay in transmission, the possible impact on the estimates.
- Recently in [10], [11], [12], [13] the most popular method called geostatistical methods is the Turning Bands method used for forecasting the performance. Another geostatistical method is sequential Gaussian simulation (SGS) [1], with the main difference between it and the TB method being the possibility of utilizing sequential neighborhoods to perform a prognosis. Because they are in an acceptable amount of these two methods was used by us Geostatisticalmethods enough climate, geology, ecology, or study agriculture developed in the traditional science of geostatistics; In addition to these, do we, for example, [2] can see-[13] where the very promising results were obtained in a study of

predicting load distribution grid, where there is usually a complex task [14] transmission network in forecast accuracy limitations although this suffered from poor prognosis is well fit for all kinds of measurements.

- In [15], the authors Lee Y.S., Tong L.I. presented the forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. This method is hybrid and based on concepts of ANN.

3. PROPOSED APPROACH FRAMEWORK AND DESIGN

3.1 Problem Definition

In this literature, we have studied the many methods those are presented for forecasting of web system performance by considering the different real time applications. In [1] we have studied the TB geostastical performance prediction method for distributed computer systems and web systems. Following two flowcharts with reference this is presented in figure 3. This system is suffered from the limitation of poor forecasting accuracy.

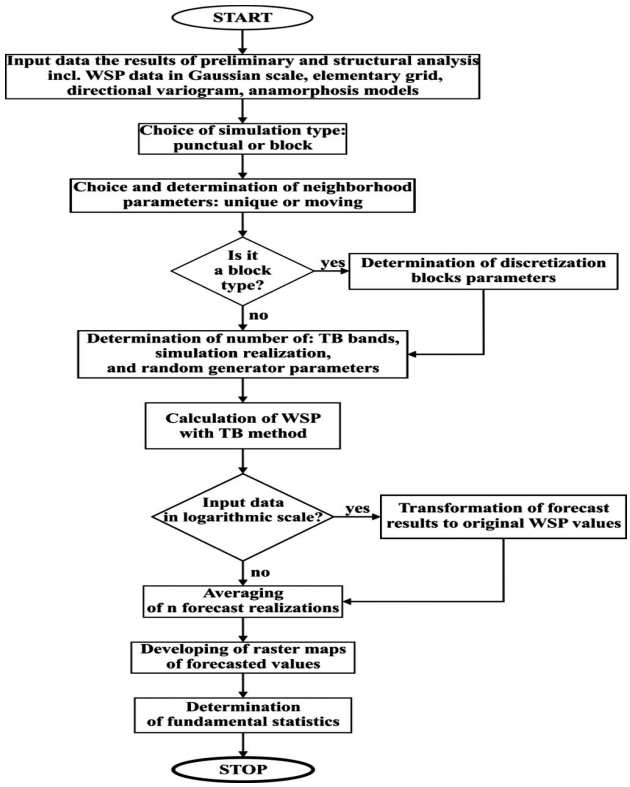

Figure 3: TB based flowchart for forecasting performance

3.2 Proposed System Architecture

In this we are presenting the new architecture which is based on two concepts first is TB which depicted in [1] and another one is FFNN which is depicted in [2]. Based on these terminologies below is proposed architecture for predicting the performance in figure 4.

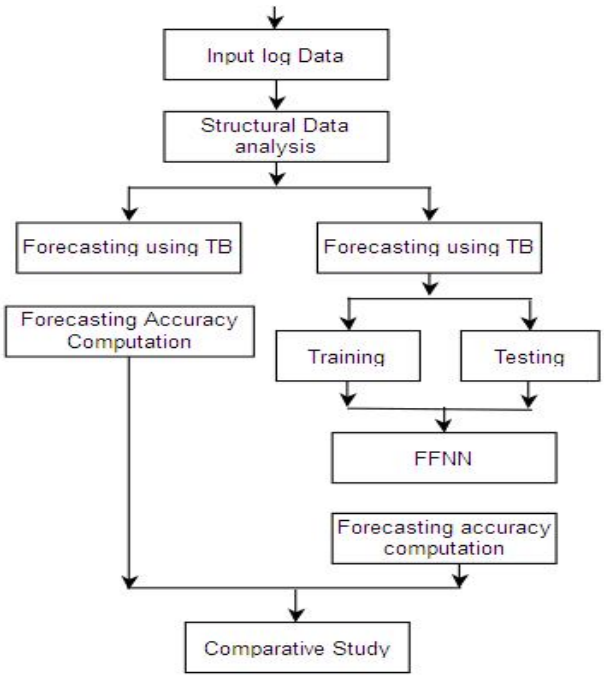

Figure 4: Proposed architecture for predicting the performance

1.TB Algorithm

1. Directions Selection.
2. Generation of standard, independent stochastic processes with covariance functions.
3. Calculate Web system Performance.

2.Proposed System Algorithm

First take inputs for proposed algorithm is a output of temporal data with some extra parameters then Training Neural Networks

- Let us invert the previous problem: Suppose that the inputs to the network are x1=1andx2=0, and is a step function.–Find values of the weights, wij, such that he output of the network y=0?
- This problem is more difficult, because there are more unknowns (weights)than known (input and output). In general, there are an infinite number of solutions.
- The process of finding a set of weights such that for given input the network produces the desired output is called training.

3.Supervised Learning

- Algorithms for training neural networks can besupervised (i.e. with a 'teacher') and unsupervised (self-organizing).
- Supervised algorithms use training set— a set of pairs (x,y) of inputs with their corresponding desired outputs.
- We may think of a training set s set of examples.
- An outline of a supervised learning algorithm

1. Initially, set all the weights wij to some random values
2. Repeat

(a)Feed the network with an input x from one of he examples in the training set

(b)Compute the network's outputf(x)

(c)Change the weightswijof the nodes

3. Until the errors(y,f(x)) is small

3.3 Mathematical Model

1. Initiation

Let S be the system that describes log file as input to system with Preprocessing of logs, patio temporal prediction, turning band, and this all gives output foretasted time to download from particular server

S= (Lf, Pf, Fn, Tb, Ft)

S= System

Lf= Log file of website

Pf= Preprocessing of log file

Fn= Feed Forward Neural Network

Tb= Turing Band

Ft= Forecasted Time to download File

2. Get website log file as input to system

Lf= (date, time ,s-ip ,cs-method ,cs-uri-stem ,cs-uri-query ,s-port, cs-username ,c-ip ,cs(User-Agent) ,sc-status ,sc-substatus ,sc-win32-status ,time-taken)

Date= date of day

Time= time of day

s-ip= source ip address

Cs-method= Post/ Get method

cs-uri-stem= URL of page

cs-uri-query= url query

s-port= http port number

cs-username= username

c-ip= client ip address

cs(user-Agent)= user agent(browser used)

sc-status= http status

sc-substatus= http substatus

sc-win32-status= http windows status

time taken= time taken to load page

3. **Preprocessing on Log File it contains**

Pt= (Fi, Cl)

Fi= Filtering of log file (here we only take necessary data as input to system)

Output->Fi= {date, time, s-ip, cs-uri-stem, c-ip, sc-status}

Cl= cleaning of data from log file

If(sc-status Ɛ { 404, 302, 500, 304})

Remove record R from log File Lf

If(cs-uri-stem Ɛ {image extensions}

Remove record R from log file Lf

4. FNN

ARFIMA models are used to model long range dependent time series. ARFIMA models were introduced by [11]. AR-FIMA model can be given by (p, d, q);

(B) $(1-B)X_t = \theta(B)e_t$, -1/2 <d <1/2

where B is the back-shift operator such that $BX_t=X_{t-1}$ and e_t is a white noise process with $E(e_t)=0$ and variance δ2. The polynomials (B)=(1-1B…….pBp) and θ(B)=(1-θ_{1B}.....θ_qBq) have order p and q respectively with all their roots outside the unit circle.[3] extended the estimation of ARFIMA models for any d > -1/2 by considering the following variation of the ARFIMA model:

(B) $(1-B)^d(1-B)X_t = \theta(B)e_t$, -1/2 <d <1/2

The integer *m* is the number of times that $_{Xt}$must be differenced to achieve stationary, and thus differencing parameter is given by d=δ+m..

5.**Turning Band Tb**

Let (θ_n, n € ℕ) be a sequence of directions , and let (Xn, n € ℕ) be a sequence of independent stochastic processes of covariance. The algorithm of the TB method takes the following form: TB Algorithm:

1. Directions Selection: $\theta_{1...}$ $\theta_n$so that 1/n $\sum_{k=}^{n}$ θ_k is weakly convergent to

θ_n=(cos(2Πun) $\sqrt{1-v^2n}$, sin(2 Πun) $\sqrt{1-v^2n}$, vn).

Un= binary expansion of each integer n=1,2,…….;

Vn= ternary expansion of each integer n=1,2……..;

2. Covariance function

$C(n)(h) = 1/n + \sum_{k=1}^{n}(C\theta k\,(< h, \theta k >))$ ~ $\theta_n = \theta_k$

3. calculate x Ɛ D

$$1/\mathrm{sqrt}(n) + \sum_{k=1}^{n}(C\theta k\,(< h, \theta k >))$$

X= forecasted download time;

4. WORK DONE

In this section we represent the input, result of practical work and environment used for implementation

4.1 Input Dataset

For this implementation, we use the dataset of log file generated from web application. This log file used for further process of forecasting accuracy computation.

4.2 Hardware and Software Used

Hardware Configuration

- Processor - Pentium IV 2.6 ghz
- Speed - 1.1 Ghz
- RAM - 512 mbdd ram
- Monitor - 15” color
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Monitor - SVGA

4.3 Software Configuration

- Operating System - Windows XP/7/8
- Programming Language - C#. Net
- Database - Sql Server 2008
- Tool - MS Visual Studio 2010.

4.4 Metrics Computed

Results are compared with previous implemented algorithm in [1]. 1. Generate temporal Data

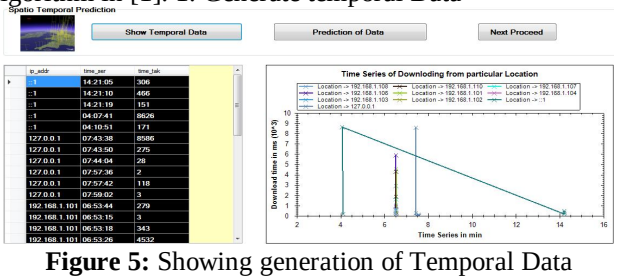

Figure 5: Showing generation of Temporal Data

Spatio Temporal prediction of data with graphical representation is shown in figure 6

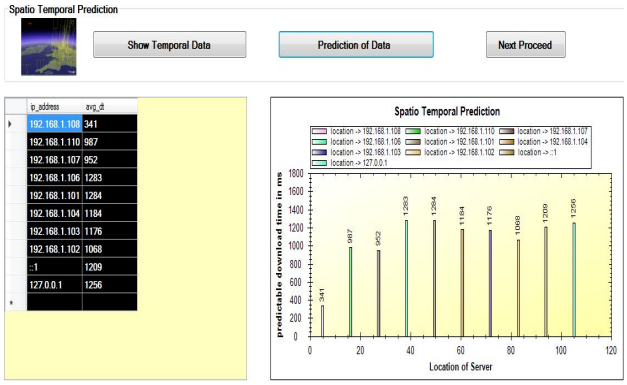

Figure 6: Spatio Temporal Prediction of data

4.5 Results of Practical Work

Result shows of existing system Using forecast values of resource download time calculated by existing system using Gaussian method and covariance function.

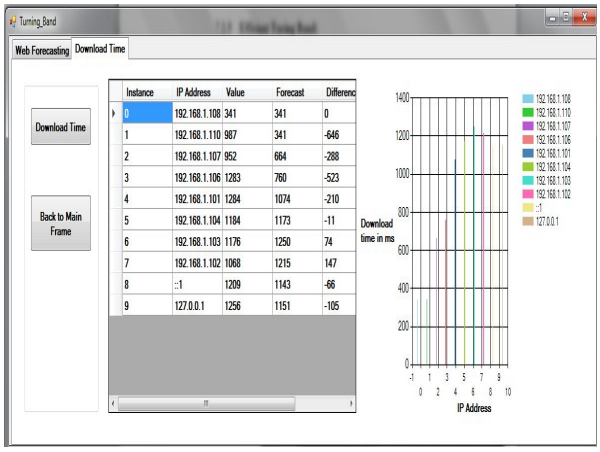

Figure 7: forecast values of resource download time using existing system(Turing Band)

Result shows of proposed system Using forecast values of resource download time calculated by proposed system using FFNN.

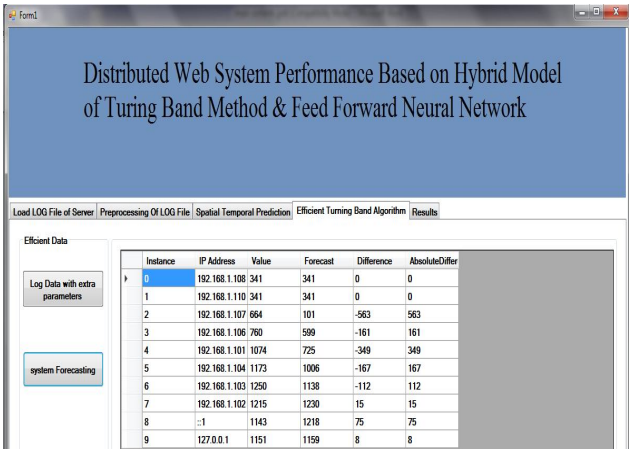

Figure 8: forecast values of resource download time using proposed system (Efficient Turing Band)

Comparative study of existing system (Turing Band Algorithm)and proposed system (Efficient Turing Band Algorithm), under calculating forecast values of resource download time by both algorithms existing and proposed also.

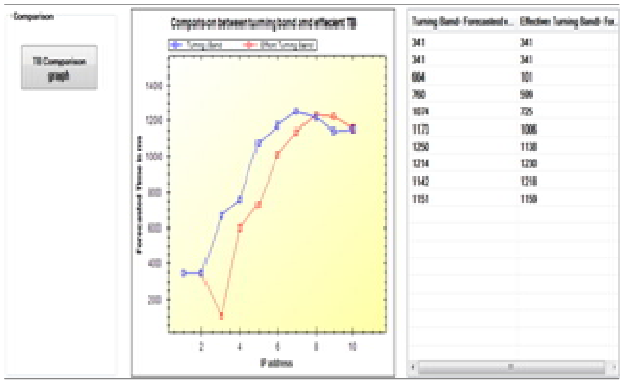

Figure9: Accuracy comparison of forecasted time to download from server.

In fig 9 x axes represents number server, y axis represent accuracy of forecasted download time, Above graph shows FNN method performs better than turning band as it gives more accurate time to download particular file from server.

5. CONCLUSION AND FUTURE WORK

In this paper we have presented the hybrid model for predicting the performance of distributed web system with aim of improving the overall forecasting accuracy. The limitation of TB method has been overcome by using this method.Finally the results which we presented in this system are based on current state of our implementation work. The TB based method is combined with FFNN for improving the performance forecasting accuracy We have achieved better performance and accuracy in forecasting of downloading time of resources from distributed web system from the log file of server using the Efficient Turing Band method. Such an application can be useful manufacturing and industrial applications over the World Wide Web (WWW) architecture. This application has been robust and gives good results even in case of many challenge like bandwidth, speed, delay etc. As a future scope of this methodology, we can improve this hybrid method and present approach which can handle the different kinds of measurements and data.

REFERENCES

[1] LeszekBorzemski and Anna Kaminska-Chuchmala, "Distributed Web Systems Performance Forecasting Using Turning Bands Method", IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 9, NO. 1, FEBRUARY 2013.

[2] CagdasHakan Aladag1, Erol Egrioglu2,*, Cem Kadilar1, "Improvement in Forecasting Accuracy Using the Hybrid Model of ARFIMA and Feed Forward Neural Network", American Journal of

Intelligent Systems 2012, 2(2): 12-17 DOI: 10.5923/j.ajis.20120202.02.

[3] M. Ulieru and S. Grobbelaar, "Engineering industrial ecosystems in a networked world," in Proc. 5th Int. IEEE Conf. Ind. Informat., Vienna, Austria, Jul. 23–27, 2007, keynote address.

[4] F. Tao, D. Zhao, Y. Hu, and Z. Zhou, "Resource service composition and its optimal-selection based on particle swarm optimization in manufacturing grid system," IEEE Trans. Ind. Inform., vol. 4, no. 4, pp. 315–327, Nov. 2008.

[5] T. Cucinotta, A. Mancina, G. F. Anastasi, G. Lipari, L. Mangeruca, R. Checcozzo, and F. Rusina, "A real-time service-oriented architecture for industrial automation," IEEE Trans. Ind. Inform., vol. 5, no. 3, pp. 267–277, Aug. 2009.

[6] L. Borzemski, K. Zatwarnicki, and A. Zatwarnicka, "Adaptive and intelligent request distribution for content delivery networks," Cyber. Syst., vol. 38, no. 8, pp. 837–857, 2007.

[7] Q. Wang and S. Gopalakrishnan, "Adapting a mainstream Internet switch architecture for multihop real-time industrial networks," IEEE Trans. Ind. Inform., vol. 6, no. 3, pp. 393–404, Aug. 2010.

[8] A. Onat, T. Naskali, E. Parlakay, and O. Mutluer, "Control over imperfect networks:Model-based predictive networked control systems," IEEE Trans. Ind. Electron., vol. 58, no. 3, pp. 905–913, Mar. 2011.

[9] L. Cristaldi, A. Ferrero, C. Muscas, S. Salicone, and R. Tinarelli, "The impact of Internet transmission on the incertainty in the electric power quality estimation by means of a distributed measurement system," IEEE Trans. Instrum. Meas., vol. 52, no. 4, pp. 1073–1078, Aug. 2003.

[10] M. Inizan, "Geostatistical validation of a marine ecosystem model using in situ data," Ecole des Mines de Paris, Tech. Rep. S-435 Centre de Geostatistique, 2002.

[11] A. Kaminska-Chuchmala and A. Wilczynski, "Application of simulation methods to spatial electric load forecasting," RynekEnergii, vol. 80, no. 1, pp. 2–9, 2009.

[12] A. Kaminska-Chuchmala and A. Wilczynski, "Spatial electric load forecasting in transmission networks with sequential Gaussian simulation method," RynekEnergii, vol. 92, no. 1, pp. 35–40, 2011.

[13] A. Kaminska-Chuchmala and A. Wilczynski, "Analysis of different methodological factors on accuracy of spatial electric load forecast performed with turning bands method," RynekEnergii, vol. 87, no. 2, pp. 54–59, 2010.

[14] T. Poplawski, "The long-term forecast of a peak power load in the electric power system—selected problems," RynekEnergii, vol. 93, no. 2, pp. 67–72, 2011.

[15] Lee Y.S., Tong L.I., "Forecasting time series using a me-thodology based on autoregressive integrated moving average and genetic programming," Knowledge-Based Systems 24(1) (2011) pp:66-72.

AUTHORS

JyotiJalindarDhayagude is student of Master of Computer Engg. , Department of Computer Engineering, JSPM's Rajarshishahu College of Engg, Tathwade, Pune 411033, India. Her research interestis Data Mining.

Dr.PradeepDeshmukh is prof. of Master of Computer Engg. , Department of Computer Engineering, JSPM's Rajarshishahu College of Engg, Tathwade, Pune 411033, India. His research interestis Data Mining and Mobile Computing