IJERT

CSECONF - 14

3rd National Conference In Current Trends In Computer Science And Engineering

14th June 2014

Organized by :

Jain University

(Department Of Computer Science and Engineering)

Conference Proceedings

Published By

International Journal of Engineering Research and Technology

(www.ijert.org)

A Survey on Load Prediction Techniques in Cloud Environment

Manjunath C R[1], Manaswini C[2], Nilasini Bangar[3]
School of Engineering and Technology
Jain University, Kanakapura, India

Abstract **--- The rapid growth of power demand from business, and Web applications has led to the emergence of cloud-oriented data centres. Load prediction is a significant cost-optimal resource allocation and energy saving approach for a cloud computing environment. Load classification before prediction is necessary to improve prediction accuracy. In this paper, a novel approach is proposed to forecast the future load for cloud-oriented data centres. First, Bayesian model is used to predict the mean load over a long-term time interval which is compared with PSR and EA-GMDH method which combines the Phase Space Reconstruction (PSR) method and the Group Method of Data Handling (GMDH) for effective prediction then Neural Network predicts the future load based on the past historical data which distinguishes itself with the presence of hidden layers followed by ssupport vector and kalmann smoother which is a multi-step-ahead CPU load prediction method based on Support Vector Regression which is very stable, i.e. its prediction error increases quite slowly as the predicted steps increase.**

Index Terms **- Cloud computing, Load prediction, Prediction accuracy.**

I. INTRODUCTION

Cloud computing is a term used to refer to a model of network computing where a program or application runs on a connected server or servers rather than on a local computing device such as a PC, tablet or smartphone. Service delivery in Cloud Computing comprises three different service models, namely Infrastructure-as -a-Service(IaaS),Platform-as-a-Service (PaaS) and Software-as-Service (SaaS). Software-as-a-Service provides complete applications to a cloud's end user. It is mainly accessed through a web portal and service oriented architectures based on web service technologies. Platform-as-a-service comprises the environment for developing and provisioning cloud applications. The principal users of this layer are developers seeking to develop and run a cloud application for a particular platform. The services on the infrastructure layer are used to access essential IT resources that are combined under the heading Infrastructure-as-a-Service (IaaS). These essential IT resources include services linked to computing resources, data storage resources, and the communications channel. Cloud computing platforms are being increasingly utilized by industry, government and academia due to their ability to deliver robust, resilient and scalable computational power. In cloud computing data centres, Giga-bit speed, or faster, networks interconnect both physical and virtual computers. These systems are dynamically provisioned based on a determination of the required computing resources requested by the end user of the cloud application. [1][3][6]

Predicting the processor availability for a new process or task in computer network systems is a basic problem arising in many important contexts. Making such predictions is not easy because of the dynamic nature of current computer systems and their workload. To ensure high scalability, flexibility, and cost effectiveness, cloud platforms need to be able to quickly plan and provide resources, which will ensure that supporting infrastructures can closely match the needs of various applications. Cloud platforms require mechanisms to continuously characterize and predict their loads.

Load prediction is a crucial issue for efficient resource utilization in a dynamic cloud computing environment based on future load prediction and an estimate of the future performance of cloud system. Effective load prediction will help administrators take appropriate actions in preventing the system suffering from traffic surge which is caused by high load. The key to accurate load prediction in cloud computing is proper modelling of the relationship between historic data and future values, and a proper understanding of cloud computing backend workloads.

II. SIGNIFICANCE OF LOAD PREDICTION

Load prediction is an estimation of demand at some future period. This architectural framework is presented in Fig. 1.[3]

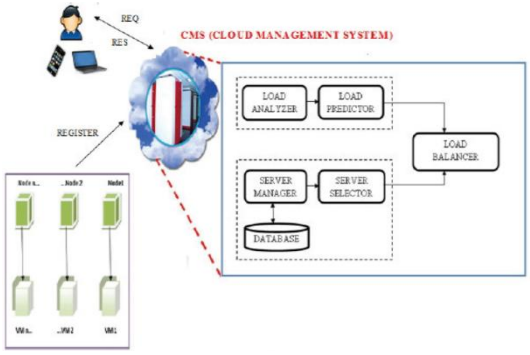


Fig. 1 Cloud architecture framework

The architecture describes how the cloud controller acts as an interface between the cloud service provider and external users which involves Load Analysis & Load Prediction, Management & Selection, Load Balancing.

In the past few years, some studies have been devoted to load prediction in cloud computing environments. This paper gives comparative study of different techniques used for load prediction. Firstly Bayesian method is discussed, which is an effective Cloud load prediction method that can accurately predict host load over a long-term period up to 16 hours in length. Prediction method based on Bayesian model is used to predict the mean load over a long-term time interval, as well as the mean load in consecutive future time intervals. It focuses on CPU. Using a Bayesian model for prediction effectively retains the important information about load fluctuation and noise. [1]

PSR and EA-GMDH is new prediction method which combines the Phase Space Reconstruction (PSR) method and the Group Method of Data Handling (GMDH) based on Evolutionary Algorithm (EA). It predicts not only the mean load in consecutive future time intervals, but also the actual load in each consecutive future time interval. PSR is an important step in local prediction methods because with a set of appropriate variables, we can reconstruct the time series. The GMDH method is a self organizing method and it has been applied to solve many prediction problems with success. [2]

Neural Network is used for load prediction, which predicts the future load based on the past historical data. It is a machine that is designed to model the way in which the brain performs a particular task. A neural network is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experimental knowledge and making it available for use. Neural network distinguishes itself by the presence of one or more hidden layers. The input signal is applied to the neurons in the second layer. The output signal of second layer is used as inputs to the third layer, and so on for the rest of the network. [3][4]

Support vector and kalmann smoother is multi-step-ahead CPU load prediction method based on Support Vector Regression which is suitable for the dynamic characteristics of applications and the complex Cloud computing environment. Kalman smoothing technology is integrated to further reduce the prediction error. It is suitable for the complex and dynamic characteristics of the Cloud computing environment. KSSVR is very stable, i.e. its prediction error increases quite slowly as the predicted steps increase. SVM has strict theory and mathematical foundation which could not lead to local optimization and dimensional disaster. [5]

III. TECHNIQUES FOR HOST LOAD PREDICTION

A. *Prediction using Bayesian Model*

Prediction method based on Bayes model is used to predict the mean load over a long-term time interval, as well as the mean load in consecutive future time intervals. Design an effective Cloud load prediction method that can accurately predict host load over a long term period up to 16 hours in length. This approach is to use a Bayesian model for prediction as it effectively retains the important information about load fluctuation and noise. Here Bayesian prediction method is evaluated using a detailed 1-month load trace of a Google data centre with thousands of machines [1].

Objective is to predict the fluctuation of host load over a long-term period, and aim is two-fold. First, at a current time point t0, predict the mean load over a single interval, starting from t0. Second, predict the mean load over consecutive time intervals. A new metric, namely exponentially segmented pattern (ESP), to characterize the host load fluctuation over some time period. For any specified prediction interval, it is into a set of consecutive segments, whose lengths increase exponentially. It predicts the mean load over each time segment. Shown in Figure 1 is an example of ESP. It denotes the total prediction interval length as s. The first segment (denoted by s1) is called baseline segment with length b, starts from the current time point t0 and ends at t0 +b. The length of each following segment (denoted by si) is $b \cdot 2^{i-2}$, where i = 2, 3, 4, · · · . For example, if b is set to 1 hour, the entire prediction interval length s could be equal to 16 (=1+1+2+4+8) hours. For each segment, predict the mean host load. The mean values are denoted by li, where i = 1, 2, 3, · · ·

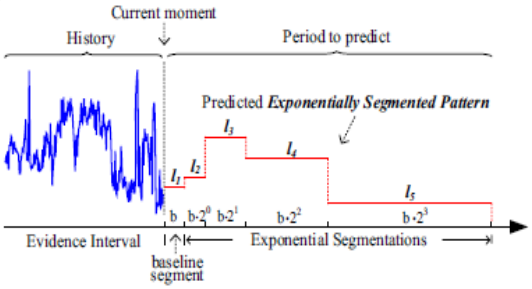


Fig. 2 Illustration of exponential segmented pattern

As illustrated above, aim is to predict the vector of load values (denoted by l), where each value represents the mean load value over a particular segment. To predict load, a predictor often uses recent load samples. The interval that encloses the recent samples used in the prediction is called evidence interval or evidence window. Given the prediction problem, one approach for prediction is to use feedback control. One could dynamically validate the prediction accuracy at runtime, adjusting the predicted values in the next interval by the error in the previous one. Then, prediction error could converge to a low level. This idea is based on the feed-back control model, which is often used in the one-step look-ahead prediction scenario.

Another approach is to use error of short-interval prediction to tune the long-term prediction. For instance, using the prediction error in a 4-hour interval may forecast the prediction error in the 8-hour interval, such that the predicted values could be tuned accordingly. However, this idea is also

inapplicable to Cloud load prediction in that short term prediction error always lags behind long-term error .

B. Neural Network Load Prediction

A neural network is a machine that is designed to model the way in which the brain performs a particular task. The network is implemented by using electronic components or is simulated in software on a digital computer.

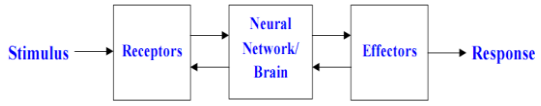


Fig. 3 Block Diagram of a Human Nervous System.

The receptors collect information from the environment. The effectors generate interactions with the environment. The flow of information/activation is represented by arrows.

Multilayer Feedforward Networks: The Feedforward neural network distinguishes itself by the presence of one or more hidden layers, whose computational nodes are correspondingly called hidden neurons. The function of hidden neuron is to intervene between the external input and the network output in some useful manner. The input signal is applied to the neurons in the second layer. The output signal of second layer is used as inputs to the third layer, and so on for the rest of the network. [3][4]

Back propagation algorithm: Multiple layers have been applied successfully to solve some difficult diverse problems by training them in a supervised manner with a highly popular algorithm known as the error back-propagation algorithm. This algorithm is based on the error-correction learning rule. Error back-propagation learning consists of two passes through the different layers of the network: a forward pass and a backward pass. In the forward pass, an input vector is applied to the nodes of the network, and its effect propagates through the network layer by layer. Finally, a set of outputs is produced as the actual response of the network. During the forward pass the weights of the networks are all fixed. During the backward pass, the weights are all adjusted in accordance with an error correction rule. The actual response of the network is subtracted from a desired response to produce an error signal. This error signal is then propagated backward through the network, against the direction of synaptic connections. The weights are adjusted to make the actual response of the network move closer to the desired response.

The load on each server is predicted for optimal load balancing. A neural network model consists of three layers with five input nodes. Fig 4 depicts the neural model. The input layer of neurons in the neural model receives five inputs from the external information source. When the network is run, each layer performs the calculation on the input and transfers the result Y_{n+1} I to the next layer.

$$Y_{n+1} = h\left[\left(\sum_{i=1}^{n} X_i w_i + b \right) / n \right] \quad \text{------ (1)}$$

The above equation (1) is used for prediction of future load based on the input value. Where Yn+1 provides the output of the current node and n is the number of nodes in the previous layer, Xi is the input of the current node from the previous layer b is the bias value and w, is the modified weight based on the mean square error and our proposed algorithm.

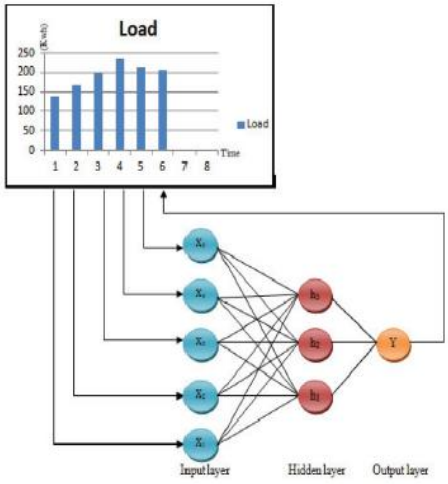


Fig. 4 Neural model

Here the neural predictor is developed and the experiment is performed to prove its highly accurate prediction. A sample load of a datacenter is analysed and given as input for the neural model.

C. Support vector and kalmann smoother

A multi-step-ahead CPU load prediction method based on Support Vector Regression which is suitable for Cloud computing environment. Kalman smoothing technology is integrated to further reduce the prediction error. Real trace data were used to verify the prediction accuracy and stability of this method. The focus of this work is on improving the CPU utilization by load prediction. KSSVR integrates SVR algorithm and Kalman smoothing technology. Furthermore, KSSVR is very stable, i.e. its prediction error increases quite slowly as the predicted steps increase. [5]

- ***Support Vector Machine:*** SVM was used for many machine learning tasks such as pattern recognition, object classification and regression analysis. It is based on the structural risk minimization principle which tries to control the model complexity as well as the upper bound of generalization risk. The principle is based on the fact that the generalization e+-rror is bounded by the sum of the empirical error and a confidence interval term that depends on the Vapnik – Chervonenkis (VC) dimension. On the contrary, traditional regression

techniques, including traditional Artificial Neural Networks (ANN) , are based on empirical risk minimization principle, which tries to minimize the training error only. Its learning process is quite complex and inefficient for modeling, and the choices of model structures and parameters are lack of strict theory. So, it may suffer from over-fitting or under-fitting with ill chosen parameters. In contrast, SVM has strict theory and mathematical foundation which could not lead to local optimization and dimensional disaster. It can achieve higher generalization performance especially for small samples set. It has a limited number of parameters to choose for modeling, and there exist fast and memory-efficient algorithms.

- ***Kalman Smoother:*** The Kalman filter has been widely used in the area of autonomous or assisted navigation. It is Kalman smoother is suitable for the Cloud application's load estimation because it was originally developed to estimate time-varying states in dynamic systems. This approach essentially uses a filtering technique to eliminate the noise of resources usage signal coming from error of measurement technique while still discovering its real main fluctuations in order to achieve a better QoS and higher resource utilization in Cloud.

D. Prediction Based on PSR and EA-GMDH

A new prediction method which combines the Phase Space Reconstruction (PSR) method and the Group Method of Data Handling (GMDH) based on Evolutionary Algorithm (EA). The proposed method could predict not only the mean load in consecutive future time intervals, but also the actual load in each consecutive future time interval. This method outperforms the other methods by more than 60% in mean load prediction, and performs well on actual load prediction over different time intervals, i.e. 0.5h to 3h. The main idea of this approach is to use PSR method and GMDH method based on evolutionary algorithm for host load prediction. PSR is an important step in local prediction methods because with a set of appropriate variables, we can reconstruct the time series. The GMDH method is a self organizing method and it has been applied to solve many prediction problems with success. [2]

The Representation of the EA-GMDH Network : To combine the EA and the GMDH network, we should first consider the representation of the EA-GMDH network. The representation of the EA-GMDH network should contain the number of input variables for each neuron, what is the best type of the polynomials for each neuron, and which input variables should be chosen for each neuron. Therefore, the chromosome for each individual should contain tree sub-chromosomes. Each sub-chromosome is represented as a string of integer

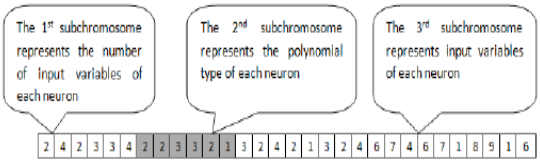


Fig. 5 The chromosome represents the EA-GMDH network

This EA-GMDH network consists of three layers, the number of neurons of each layer are 3, 2 and 1. The number of input variables of each neuron ranges from 2 to 4, and the type of polynomials ranges from 1 to 3.

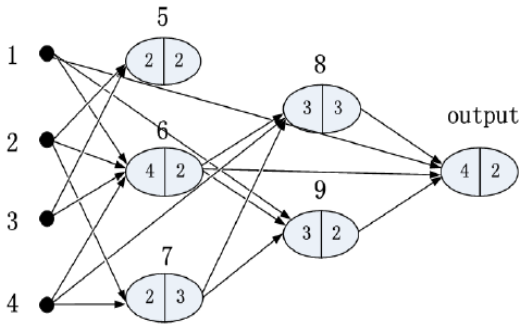


Fig. 6 The structure of the EA-GMDH network

The training set is used to calculate the coefficients of each neuron of the model. The validation set is used to evaluate each individual in each generation according to the fitness function. And the prediction set is used to estimate the performance of the model. The output of this proposed is a vector of the host load, which will not generate cumulative errors regardless of the step length, as the current predict value has nothing to do with the last predict value. We quantified the performance of actual load prediction with mean squared error (MSE).

IV. CONCLUSION

This paper summarizes the classification of load prediction methods and its impact on cloud environment. Some of the methods as discussed in the below table mainly focuses on host load prediction. We conclude that PSR & EA-GMDH outperforms other algorithms for dynamic cloud load prediction.[6] [2] It shows very good performance in long term load prediction with high performance accuracy and least error rate (MSE).

TECHNIQUE	DESCRIPTION	PREDICTION DURATION	MSE	SUMMARY/FINDINGS
Bayesian	Predicts the mean load over a longtime period as well as consecutive future time intervals.	Upto 16 hours	Lesser error rate	Retains information about load fluctuation and noise using ESP
GMDH & PSR	Predicts the mean load over a long time period as well as the mean load in the consecutive future time intervals.	0.5 to 3 hours	Least error rate	PSR: a set of appropriate variables, can reconstruct the time series. GMDH: self organizing.
Neural Network	Predicts the future load based on the past historical data.	1 second to 90 second	Medium error rate	Suitable for short term period
KSSVR	multi-step-ahead CPU load prediction method based on Support Vector Regression and Kalmann smoother methods.	1 to 2 hours	Highest error rate	Suitable for complex and dynamic characteristics of cloud environment. It is stable as the prediction errors increases quite slowly with prediction stages.

Table. 1 Comparision of Different Load Prediction Technique

REFERNCES :

[1] Sheng Di, Derrick Kondo1, Walfredo Cirne "Host Load Prediction in Google Compute Cloud with a Bayesian Model" France, 2Google Inc., USA, 2012 IEEE.

[2] Qiangpeng Yang, Chenglei Peng, Yao Yu, He Zhao, Yu Zhou, Ziqiang Wang, Sidan Du "Host Load Prediction Based on PSR and EA-GMDH for Cloud Computing System" 2013 IEEE Third International Conference on Cloud and Green Computing.

[3] Iniya, Venkatalakshmi, Ranjithflalakrishnan "Neural Load Prediction Technique for Power Optimization in Cloud Management System" Proceedings of2013 IEEE Conference on Information and Communication Technologies (IC2013).

[4] John J. Prevost, KranthiManoj Nagothu, Brian Kelley and Mo Jamshidi, Electrical and Computer Engineering "Prediction of Cloud Data Center Networks Loads Using Stochastic and Neural Models" proceedings of 6th international conference, 2011 IEEE.

[5] Rongdong Hu, Jingfei Jiang, Guangming Liu, Lixin Wang "CPU Load Prediction Using Support Vector Regression and Kalman Smoother for Cloud" 2013 IEEE 33rd International Conference.

[6] Da-yu XU, Shan-lin YANG, Ren-ping LIU "load prediction in cloud-oriented data centers" on Distributed Computing Systems Workshops, Journal of Zhejiang University 2013-14.

Benchmark Approach for Regional Climate Model in HPC Platforms

Indu B [#1], K C Gouda [*2], Nirmala H [#3]

[#]Department of Computer Science and Engineering, SJB Institute of Technology (VTU)
BGS Health & Education City, Kengeri, Bangalore-60, India

[*]CSIR Centre for Mathematical Modelling and Computer Simulation (C-MMACS)
Wind Tunnel Road, Bangalore-37, India

Abstract **— Now a days several models are available for the regional climate studies like Weather Research and Forecasting (WRF) Model, MM5 model, etc. which are designed to serve both atmospheric research and operational forecasting needs. RegCM is a Regional Climate Model Regional climate models (RCMs) are widely used tools to produce high resolution climate simulations at regional scales. In the growing age of HPC and Modelling, Satellite data its very important to use the modelling approaches to understand the weather and climate at regional scale (with very high resolution). The model has the ability to produce simulations reflecting either real data or idealized atmospheric conditions. It provides a flexible and computationally efficient platform for operational forecasting, The model also offers advances in physics, data assimilation and numerics. It can be more precisely defined as knowledge based system development that will provide researchers and atmospheric scientist to choose and designed a more advanced and doctored approach for predicting weather and configuring the system. In this paper an approach of a benchmark of RegCM is studied using the HPC system.**

Keywords— RegCM, Model Dynamics, Model Benchmarking, HPC.

I. INTRODUCTION

Weather is the day-to-day state of the atmosphere in a region, and its short-term variation whereas Climate is defined as statistical weather information that describes the variation of weather at a given place for a specified interval. Both are used interchangeably sometimes but differ in their measure of time, and trends that affect them. To understand the behavior of the atmosphere and predict the hereafter variation various models were used: global or regional. Various organizations and research institutes carryout the research activities to simplify the process with high end technologies and framework by developing sophisticated climate models. Climate models describe several compartments of the climate system, as for instance, the atmosphere, the oceans, the cry sphere, the surface hydrology, the vegetation, or cycles of matter.

The meso-scale weather prediction Model is designed to serve both atmospheric research and operational forecasting needs it is a next-generation meso-scale numerical weather prediction system. For the purpose of parallel computation and system extensibility it features two dynamical cores, Several applications including wide range of meteorological applications across scales ranging from meters to thousands of kilometers. The model has the ability to produce simulations reflecting either real data or idealized atmospheric conditions. It provides a flexible and computationally efficient platform for operational forecasting, The model also offers advances in physics, data assimilation and numerics. The objective of this study is to make aware of various sensitive features and understand the model from insight. Model physics includes understanding the various laws and factors that governs the interaction between model components and the matter of the atmosphere. It includes various convection schemes, boundary layer, radiative transfer, land surface processes etc.

The applications of the weather and climate models include:

- To predict the future atmospheric changes
- To prevent catastrophic storms, cyclones etc
- To understand the rainfall, vegetation, land usage of particular region.
- To help in initiating plan for extreme weather events
- To support city planners and architects.

II. LITERATURE REVIEW

The Regional Climate Model system RegCM ,which is maintained in the Earth System Physics (ESP)which is a section of the ICTP, It was developed originally at the National Center for Atmospheric Research (NCAR). This model provides a flexible, portable and easy to use platform. It can be applied to a large region of the World, with the grid spacing of up to about 10,000mts (hydrostatic limit), and for a wide range of studies, from process studies to paleo-climate and future climate simulation. The history of numerical weather prediction began in the 1920s through the efforts of Lewis Fry Richardson who utilized procedures developed by Vilhelm Bjerknes. Prior to Norwegian scientist Bjerknes, the chapter of meteorology was opened by Cleveland Abbe [1]. It was not until the advent of the computer and computer

simulation that computation time was reduced to less than the forecast period. ENIAC initiated and created the first computer forecasting system in 1950, and more powerful computers later increased the size of initial datasets and included more complicated versions of the equations of motion. In 1966,West Germany and the United States began producing operational forecasts based on primitive-equation models, followed by the United Kingdom in 1972 and Australia in 1977 [1]. The development of global forecasting models led to the first climate models. The development of limited area (regional) models facilitated advances in forecasting the tracks of tropical cyclone as well as air quality in the 1970s and 1980s.

A. Overview of HPC

High Performance Computing (HPC) is the generic name for the most powerful system available at the front line of current processing and simulating capacity, particularly in terms of speed of calculation. The term "SUPERCOMPUTER" is used to denote such class of system that can advance knowledge and generate insight that would not be otherwise be possible or that could not be captured in time to be actionable. They are the indispensable tools for solving the most challenging and complex scientific and engineering problems including the simulation and modeling of physical phenomena. As the technology advances to a new era the core component of any computation i.e., data is not stable and to compete with such petabytes of data's modest computing system is required. However many fast processing systems were developed, yet they are dawdled by scaling, timeliness, architectural design and ability to address important issues. Thanks to the advancement in petascale computing technologies that will overcome the processing and performance constrain of computing resources. One motivation of such computing is to aggregate the power of multiple system in to single system to study the high end calculation intensive task that cannot be possible with single core system. To achieve this goal, proper understanding of system tools, software and underlying hardware is essential.

The development can be traced back to 1960 with the initiation of first supercomputer CDC 6600 by Seymour Roger Cray at Control Data Corporation. With time their developed more such system that can exploit the processing speed and performance, few with thousands of processors and others more than that. Though the market was flourished by European design yet, India was no way out from the challenge, and the development was marked by India's first supercomputer Param 8000 built in 1990 by Center for Development of Advanced Computing (CDAC).

B.Applications of HPC in Weather Forecasting

Supercomputers are the body builders of the computing world. They boost the computing power and cost millions of dollars. The machine can be used in both scientific and business applications, mostly they are used to tackle scientific calculation. It finds its uses from uncovering the origins of the universe to delving into the patterns of protein folding that make life possible. The few classes of applications are listed below. Several studies and projects are being carried out in NASA, MRI Japan, Meteo-France, WMO, NCEP, NCAR, ECMWF in internationaly and MoES, IITM, CSIR C-MMACS, NCMRWF, IMD, ISRO etc in India regarding the Weather prediction and forecasting system for the accurate and advanced prediction of monsoon, disasters like Cyclone, Extreme weather events etc. to avoid the losses of live and money

III. System Model Description

All models share the same objectives: process studies, diagnosis and forecasting. The four major components of modern coupled climate models which includes atmosphere, ocean, land surface, and sea. The development of each component raises important questions as to what is the road map to guide the development of the model architecture and how the physical processes are represented in models. Climate modeling has been steadily improving over the past several decades, but the pace has been uneven because the goal of simulation has several important aspects of the climate system that present severe challenges.

The design of this system consists of two different inputs called as initial and boundary condition. Initial conditions generally taken from different satellites on timely basis and Boundary condition generally taken from the historic or climatological data as well as validate terrain and ungrib dataset. It describes the detailed model configuration needed for the RegCM model. On a long run, statistics can be accumulated that give information on the performance of a particular model or forecast system. On the other hand In climate change simulations, the models can be used for projections of possible future changes over time scales of many decades and for which there are no precise past details

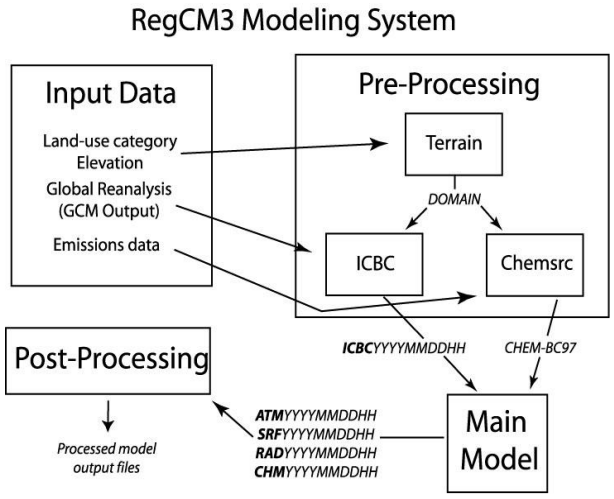


Fig 1 RegCM Flow of Work

The boundary conditions derived from Global Climate Model output can be assimilated into the RegCM3. The Boundary conditions for Regional Climate Models are derived by preprocessing saved GCM fields into a standard format that can be read into the model. The RegCM4 model is a hydrostatic, sigma vertical coordinate, compressible model that runs on different grid structure with various convection scheme in which the atmospheric components like wind, rainfall, temperature etc and thermodynamic variables are staggered horizontally. The use of various scheme help the model customized to different weather and climate parameters like precipitations, sea surface temperature, wind etc.

IV Results and Discussions

The model performance and sensitivity to model physics options are studied using Weather Research and Forecasting model over different region in India for surface and upper air meteorological parameters in different seasons like summer and winter seasons. Generally, the combination of Pleim-Xiu land surface model. The utilization of the Pleim surface layer scheme, and Asymmetric Convective Model will produce the better estimates of temperature and relative humidity for a region wise variations. Several case studies were undertaken in the CMMACS HPC environment for the simulation of weather and climate parameters using the RegCM climate model. Firstly the model is installed and benchmarked in three machines, a comparative evaluation and optimization resources are studied. In the present case the monsoon rainfall is simulated and compared with the IMD observed rainfall climatology. A comparative evaluation of the model execution time on different machines for simulation of 1 month duration has been studied. It is inferred that Altix machine with 12 processor in parallel run is the efficient and optimized.

The main part of the work is to understand the capabilities of the regional climate model to capture the weather and climate parameters accurately, so to study this several sensitivity studies were carried out and it is found that in general, model overestimates and underestimates pre-monsoon and monsoon rainfall, respectively. For sub-annual scale (March-November) and averages from 4 years (1991, 1994, 1996 and 1999), it is found that estimated rainfall by Model and observation is 2432.18 and 2398.30 mm, respectively. More research work is necessary on other model options and other parameters for long term data analysis. Once the option is settled then the RegCM outputs may be useful in real time and high resolution rainfall forecasting over India.

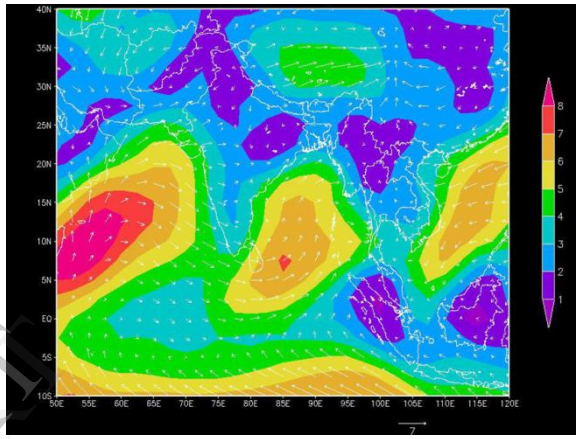


Fig 2: Plotted wind vectors for zonal and meridional components

Conclusions

The model will emphasize on the evaluation of the RegCM in the high performance computing (HPC) and cloud computing environment. Basically the model will be tested in the computing platform, how well the model performs to predict and forecast the weather and climate. Finally the evaluation of high resolution weather model for the climate studies at regional scale will be carried out with different case studies like prediction of a heavy rainfall event or cyclone etc. This model will emphasize on predicting the weather and climate at very high resolution ie at about 25 sq km.

Future Enhancement

This study can further be enhanced to study the different weather phenomena like monsoon rainfall forecasting, cyclone simulation, extreme rainfall events and extreme cold or hot days simulation etc., which in turn can be used by other sectors like agriculture, health, water and disaster sector etc. for a better forecasting assessment and mitigation of the disasters.

ACKNOWLEDGEMENT

The first author Indu B acknowledges to SPARK program of CSIR C-MMACS, HOD Computer Science and Engineering & Principal, SJB Institute of Technology, Bangalore for providing the necessary infrastructures to carry out the work.

REFERENCES

Anthes RA (1977) A cumulus parametrization scheme utilizing a one-dimensional cloud model. Mon Weather Rev 105: 1423–1438

Anthes RA, Hsie EY, Kuo YH (1987) Description of the PennState/NCAR Mesoscale Model Version 4 (MM4). Nationa Center for Atmospheric Research Tech Note tn-282+ str, ncar, boulder, coz.

Grell GA (1993) Prognostic evaluation of assumptions used by cumulus parameterization. Mon Weather Rev 121: 764–787.

Giorgi F, Marinucci MR, Bates G (1993) Development of a second generation regional climate model (RegCM2). I. Boundary layer and radiative transfer processes. Mon Weather Rev 121: 2794–2813

Emanuel KA (1991) A scheme for representing cumulus convection in large-scale models. JAtmos Sci 48: 2313–2335

Emanuel KA, Zivkovic-RothmanM(1999) Development and evaluation of a convection scheme for use in climate models. J AtmosSci 56: 1766–1782.

Pal JS, Small E, Eltahir E (2000) Simulation of regional-scale water and energy budgets: representation of subgrid cloud and precipitation processes within RegCM. J Geophys Res 105: 29579–29594

Kiehl J, Hack J, Bonan G, Boville B, Breigleb B, Williamson D, Rasch P (1996) Description of the NCAR Community Climate Model (CCM3). National Center for Atmospheric Research Tech Note NCAR/TN-420+STR, NCAR, Boulder, CO.

Peter Lynch, "The Origin of Computer Weather Prediction and Climate modeling", *School of Mathematical Sciences, University College Dublin, Journal of Computational Physics, 19th March 2007.*

Bibrak Qamar, Jahanzeb Maqbool, "Implementations and Evaluation of Scientific simulations on High Performance computing Architectures", School of Electrical Engineering & Computer Science and NUST, Pakistan, 2011.

Kit K. Szeto, "An Overview of Atmospheric Model", *Meteorological Service of Canada, MAGS Workshop, 5-6 Sepember 2002.*

David C Bader, Issac M Held, "Climate Models- An assessment of Strengths and Limitations", *US Climate Change Science Program, July 2008.*

N .T. Karonis, B toonen, et al, "MPICHG2: A Grid enabled Implementation of the Messaging passing Interface", *Journal of Parallel and distributed computing (JPDC). Vol.63(5), PP:551—563.2003.*

Ehsan Mousavi Khaneghah " Evaluating the Effect of Interprocess Communication Efficiency on High Performance Distributed Scientific Computing", *IEEE International Conference on Embedded and Ubiquitous Computing, DOI: 10.1109/EUC.208.11.*

Tim Killeen and Mehmet Celenk, "Reducing Interprocess communication Overhead through Resgister Windows", Dept. of ECE, Ohio University, Athens, IEEE ,1995.

Carla Osthoff, Claudio Schepke "I/O Perfromance on Multicore ClusterswithAtmosphericModelEnvironment",*http://www.gppd.inf.ufrgs.br//atmosferamassiva*

J A Smith, "hpsgprof: A New Proling Tool for Large{Scale Parallel Scientific Codes", *High Performance Systems Group, Department of Computer Science, University of War-wick, Coventry, CV4 7AL jas@dcs.warwick.ac.uk*

Michael Collette, Bob Corey, and John Johnson, "High Performance Tools & Technologies", *Computing Applications & Research Dept., Lawrence Livermore National Laboratory, December 2004.*

F. Giorgi, G. Bates, and S. Hostetler, Towards the simulation of the effects of the great lakes on regional climate, Monthly Weather Review, 121, (1993) 1373-1387.

F. Giorgi, S. Hostetler, and C. Shields Brodeur, Analysis of the Surface Hydrology in a Regional Climate Model, Quarterly Journal of the Royal Meteorological Society, 120, (1994) 161-184.

Defragmentation Based Dynamic Storage Allocation in Cloud

Krishna Prasad .R[1], *Madhu B.R.*[2]
[1]M.Tech-CSE, SET, JAIN University,
[2]Asst Prof, Dept of CSE, SET, JAIN University

Abstract - Cloud computing is an emerging computing paradigm that supply's users everything as a service. Cloud computing arise as an efficient way toallocate resources like storage, memory and processor for execution of task and services for a geographically dispersed providers from different organizations. Focus is on resource management problem is transformed to resource virtualization and allocation instead of job decomposition and scheduling. Public clouds sell capacity in the form of pre-defined virtual machine (VM) configurations to their users. This force's the users to buy the VM configuration based on the peak usage.This reduces the value proposition of moving to apublic cloud as compared to doing consolidation in a privatevirtualized datacentre. Ideally we would like the cloudtenants to buy capacity in bulk and benefit from statisticalmultiplexing among workloads. This requires dynamic allocationof bulk capacity among VMs of a user that may berunning on different servers across different datacentres.In this paper, we have used a Basic Base plus Proportional Excess model which is capable of adaptively adjusting storage space in which owner plays important role in storage allocation. We see several remarkable observations of the relationships between users and providers and owners.

IINTRODUCTION

Cloud computing is a technology which uses internet to work. Data and Applications are stored and maintained using Central remote servers in cloud. Applications can be used in Cloud computing without any installation of software or any specific hardware. The users can access the Internet and send messages anywhere in the world. Centralized storage, memory, processing and bandwidth are more efficient incomputation thatare allowed in Cloud computing. There are several cloud providers cloud services. Most Providers offer their services at fixed price and storage area, such Amazon EC2 and Windows Azure.

Prominent industry players like Amazon, Google, Network.com and Salesforce already provide Grid-based services on demand, but often use static pricing models that do not reflect the dynamics of the market supply and demand, such as pay-per-use or subscriptions for static resource configurations.

The goal of this paper is to design a defragmentation based system for the cloud instance market where multiple cloud users and providers respectively buy and sell instances to run and host cloud-based Internet applications.

The proposed Defragmentation based system assists the cloud users (i.e., both cloud users and providers) to decide how providers allocate their storage to which users in order to improve both cost and resource efficiency. In the proposed defragmentation based system users can bid for the resources in order to buy their needed resources and if required the user can upgrade using the defragmentation technique. The contribution of the paper can be summarised as

- The approach ofdefragmentation based mechanism in the cloud market where the user gets some extra storage space with the allocatedresource.
- We propose defragmentation based mechanism for single provider and multi users and can be extended to multi provider.

IIRELATED WORK

A. Resource Allocation

The importance of resource provisioning has been well discussed in various fields such aswireless networks, energy industries, and advertisements, which have proposed the allocation and pricing model of resources (e.g., wireless channels [1], [2], electricity [3], [4], and advertisements) to improve the resource utilization and efficiency. We focus on instances in clouds and consider an instance market where computing resources (e.g., bandwidth, CPU time and memory space) are traded as instances.

For the resource allocation, there exist several techniques such as game theory finding an equilibrium solution among players [5],stochastic programing considering uncertainty [6]; and bio-inspired mechanisms (e.g., genetic algorithm that seeks a Pareto solution of a multiobjective problem [7] and Ant colony that provides a heuristic solution of a complex problem[8]). We apply the auction theory to design the cloud market and formulate its resource allocation.

B. Auction based mechanism

Auction-based mechanisms have been proposed in various fields such as wireless networksand cloud computing in order to investigate how participants (or nodes) behave in a competition for resources; and different classes of auctions such as sequential second price auction [9], Vickrey auction [10], double auction [11], and combinatorial auction [12] have been considered in the design of the mechanisms.

C. Bidding strategies

Technology trends like Grid and Cloud allow resource providers to organize their resources moreefficiently and to offer them on demand. Prominent industry players like Amazon, Google, Network.com and Salesforce already provide Grid-based services on demand, but often use static pricing models that do not reflect the dynamics of the market supply and demand, such as pay-per-use or subscriptions for static resource configurations. Although there are business models that successfully promote consumption of these kinds of distributed services, incentives for using the services and simplified means of accessing them, along with legal regulation, are still insufficient.

Requirements for bidding strategies

Bidding strategies must allow the automation of the bidding process of offering and requesting computational services.

Bidding strategies must define and implement behaviours that incorporate the desired goals (maximize profit, minimize completion time), take actions (generate bid) and calculate the effects (payoff) of the actions.

Bidding strategies should adapt to market dynamics like price fluctuations, changing conditions and QoSof the provided resources. Learningmechanisms will be adopted to aggregate information pertaining to past decisions,actions and available market information in order to “fine tune” the bid generation processes.

Bidding strategies should be flexible i.e. converge to approximately optimalbids in multiple market mechanisms.

D. Windows azure architecture

Windows Azure is an Internet-scale computing and services platform hosted in Microsoft data canters. It includes a number of features with corresponding developer services which can be used individually or together.The Windows Azure SDKs for .NET, Node.js, Java, and PHP provide common tools and resources that you use to package, test and deploy your application. The Windows Azure SDK for .NET includes the Windows Azure Tools for Microsoft Visual Studio, which extends Visual Studio to enable the creation, building, packaging, running, and debugging of scalable web applications and services on Windows Azure.

The Windows Azure Management Portal provides access to Cloud Service (hosted service) deployment and management tasks as well as at-a-glance status information that lets you know the overall health of your deployments and accounts.

The Management Portal organizes the components of your Windows Azure deployments with constantly refreshed information that is easy to discover and understand. There are two portals available at this time. The New portal and the older Silverlight based portal.

Fig 1.azure architecture

III. DEFRAGMENTATION BASED MODEL

In this paper cloud environment is modelled as 3 parts: cloud providers, cloud users and the broker. Users submit the type of resource to the broker with respective bid price. The broker plays the main role. The advantage of using broker is, broker will have the complete trading information of the auction. The broker does 2 jobs. First calculate the highest bidder for the respective type of resource and second to allocate the resource to highest bidder.
And if required the defragmentation is done on the storage that is allocated.

Algorithm

Basic Base plus Proportional Excess:
Start
Input: VM settings (v); C: Capacity to allocate.
Result: Allocation Computed for VM.
Variables: i=1…n, a, b, u, v
For each i=1…n do
/*Allocate each VM its Lower Bound
a(i)=v(i)
b(i) = u(i)-v(i)
/* E is the remaining capacity
E= C – Sum(v(i)).
For each I = 1…n do
the maximum number of VMs user may need before VM creation.

a(i)=a(i) + ((b(i) * E)) / Sum(b(i))).
End

IV.COMMUNICATION AMONG ENTITIES

Cloud users register themselves with the broker. Broker provides database level match-making services for mapping user requests to suitable cloud providers. Broker acting on behalf of users consult the cloud providers about the list of cloud resource available to bid. Cloud provider initially fixes the cost to resource according to which the user has to bid. Messages from brokers to users may require a conformation, about the execution of the action.

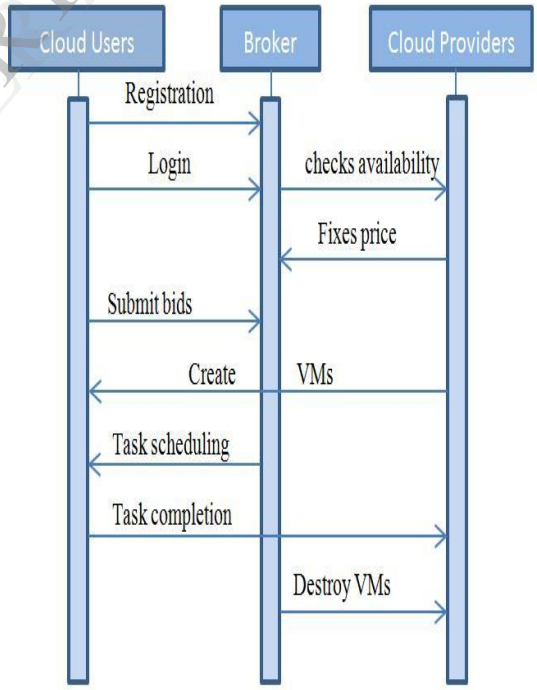


Fig2:communication data flow

V. EXPERIMENTAL SETUP FOR SINGLE PROVIDER

We use azure cloud for our single provider which provides VMs to our users. We create a broker website which acts as web hosting where the users submit their bids to the broker. We create the web role which simply acts as front end web applications. User interacts only with the web role. Broker decides to which user the VM should be allocated. Broker interacts with the provider worker role by creating the input endpoint to the worker role in the service definition file when deploying the role. Now the particular user can have full access and control over the VM, so the user can do anything he want and configure as he like. If the user requires more space he can raise a request to the owner where the owner will apply the defrag method on the leftover spaces such the user gets the required space at extra cost and SLA.

VI. RESULTS

Fig4. Storage allocated to user

The point of the diagram below is to think about

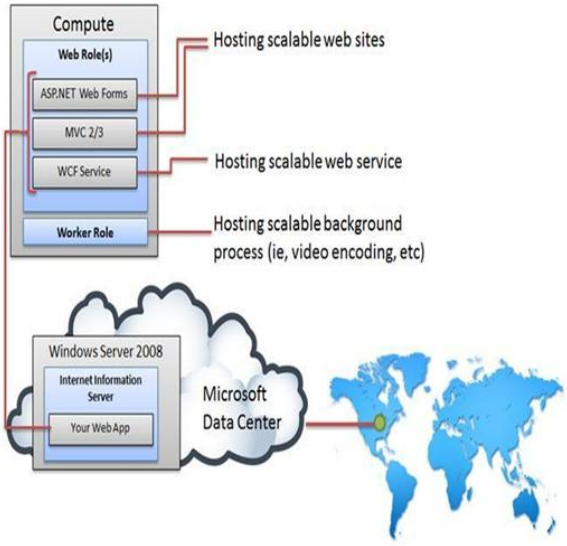


Fig3. Hosting your application

Fig5. Defragmentation done after the initial allocation is done.

VII. CONCLUSIONS AND FUTURE WORKS

In this paper we see the defrag model with the broker model important role for resource allocation with single provider. We also see the communication between the entities occurring during resource allocation. In future it would be interesting to work on multi provider and multi user to adaptively adjust resource price and storage space to see the advantages of defrag model rather on fixed price and space model for resource allocation.

REFERENCES:

[1] Linux containers (LXC) overview document. http://lxc.sourceforge.net/lxc.html.
[2] Solaris Resource Management.http://docs.sun.com/app/docs/doc/817-1592.
[3] B. Agrawal, L. Spracklen, S. Satnur, and R.Bidarkar. Vmware view 5.0 performance and best practices. 2011. http://www.vmware.com/files/pdf/view/VMware-View-Performance- Study-Best-Practices-Technical-White-Paper.pdf.
[4] D. Ardagna, M. Trubian, and L. Zhang. SLA based resource allocation policies in autonomic environments. J. Parallel Distrib. Comput., 67(3):259–270, 2007.
[5] G. Banga, P. Druschel, and J. C. Mogul. Resource containers: a new facility for resource management in server systems. In OSDI '99.
[6] H. Benjamin and et. al. Mesos: a platform for fine-grained resource sharing in the data center. In NSDI'11.
[7] G. Casale, N. Mi, L. Cherkasova, and E. Smirni.How to parameterize models with bursty workloads. SIGMETRICS Perform. Eval. Rev., 36 (2):38–44, 2008.
[8] L. Cherkasova and J. A. Rolia. R-opus: A composite framework for application performability and qos in shared resource pools. In DSN, pages 526–535, 2006.
[9] D. Gmach, J. Rolia, L. Cherkasova, and A. Kemper. Capacity management and demand prediction for next generation data centers. In ICWS, pages 43–50, 2007.
[10] D. Gmach, J. Rolia, L. Cherkasova, G. Belrose, T. Turicchi, and A. Kemper. An integrated approach to resource pool management: Policies, efficiency and quality metrics. In DSN, pages 326–335, 2008.
[11] D. Gmach, J. Rolia, and L. Cherkasova. Satisfying service level objectices in a self-managing resource pool. In SASO, 2009.
[12] D. Gmach, J. Rolia, and L. Cherkasova. Selling t-shirts and time shares in the cloud. In CCGRID, pages 539–546, 2012.
[13] Chih-Wei Tsai, Zsehong Tsai" Bid-Proportional Auction for Resource Allocationin Capacity-constrained Clouds" Graduate Institute of Communication Engineering, National Taiwan University,2012.

Design and Development of Cut Node Based Routing Protocol for Delay Tolerant Networks

Guru H. G.
M.Tech 4th Semester,
B.I.T, Bangalore.

Jyothi D.G.
Associate Professor,
Dept. Computer Science & Eng,
B.I.T, Bangalore.

Shobha Y.
Associate Professor,
Dept. Computer Science & Eng,
B.I.T, Bangalore.

Abstract: Routing in delay tolerant networks (DTNs) is a challenging problem in networking research. Existing DTN routing solutions have used many approaches to increase the success rate of message delivery, such as meeting probabilities between nodes, packet replication and flooding. One important feature of these protocols is using local connection information to find the "best" path with high likelihood to deliver a packet. From a global view, a general disconnected network can have many small instantaneously clustered mobile nodes. Mobility allows nodes carrying messages to deliver them to other clusters. Selecting appropriate nodes to carry and deliver messages becomes important in order to reduce message delay and overhead. The proposed protocol tackles this issue by utilizing cut nodes among a local sub-graph formed by including all neighbors of two "meeting" nodes. Cut nodes are the cut vertices of this local sub-graph, and by definition are the nodes, whose removal will disconnect the graph. Thus, these cut nodes are more likely to be able to deliver messages outside the local cluster. Packets will be buffered in these nodes and forwarded to other cut nodes when they meet. The process repeats until messages reach their destinations. The simulation results show that the proposed algorithm performs better than related protocols in terms of delivery rate and efficiency.

Keywords—DTN, MANET, Routing Protocol, Connectivity

I INTRODUCTION

Mobile network techniques allow users to communicate in situations that were not present in traditional wired Internet. Specifically, networking opportunities extend to scenarios where connections to other nodes or Internet are intermittent. Delay tolerant networking architecture (DTN) has been studied for these scenarios. In addition, a mobile ad hoc network (MANET) can turn to bearing such links when the number of nodes in the network becomes sparse. These connection outages create several disconnected partitions of the connected network. Thus, a hybrid network with DTN and MANET can be built to assist message delivery. Many applications and projects are built on this hybrid network approach such as ocean sensor networks [1], and vehicular networks [2, 3].

Traditional MANET routing protocols do not apply when the entire network is not fully connected. Instead, many DTN routing protocols are proposed recently. The main challenge of routing in the hybrid delay tolerant MANET is the uncertainty about network conditions. In the rest of the paper, we will refer to this hybrid network simply as DTN.

Researchers have proposed several approaches on how to find a high probability path to deliver the message with extremely limited local information. Those approaches include estimating the likelihood of nodes meeting by using different mechanisms and packet replication. The probabilistic approach faces the high failure rate of delivering messages since it only keeps one copy of a message along the network, while packet replication suffers the high overhead of storing and forwarding multiple copies of messages. In this paper, we propose a mixture model of routing in the hybrid network.

An observation on the character of the network in question is that it can be separated to several subnets which have high local connectivity. In those subnets, some nodes contact with other subnets intermittently. Based on this observation, we call such nodes as cut nodes in DTN. The cut nodes have higher probability to deliver the message outside the local strongly connected sub-graph. Unlike the computation of cut vertices in graph theory, the cut nodes are locally computed without the knowledge of the entire graph. The cut nodes are discovered when two nodes meet each other and exchange the routing information stored in their routing table. They may not be actual cut vertices because their removal does not necessarily disconnect the entire graph, but instead their removal disconnects the graph in a local sense.

The major contributions of the paper are: (1) We propose a new routing protocol combining forwarding and replication routing. (2) Our protocol requires relatively little overhead from computing and storing information: since each node only computes the connectivity of a local sub-graph and forward packets to the nodes which are identified as "cut nodes". (3) Our protocol provides high message delivery rates: most possible paths from source to destination are covered by this protocol.

II RELATED WORK

A. Routing Protocols in DTN: Routing in DTN has drawn a lot of interest because traditional routing approaches (e.g., routing in MANETs) cannot apply directly. People have recognized that though a path from a source to a destination does not exist at the instantaneous moment, such a path may occur over a time period. However, packets are not guaranteed to be delivered if such a path does not occur. DTN routing protocols have targeted at many application scenarios. Among them, a few have utilized real world contact patterns. Here we briefly discuss routing protocols that are close to our approach.

The simplest DTN routing protocol is epidemic flooding [4]. In this case, packets are replicated to every node which can possibly deliver the packets. Obviously it will waste resources and degrade performance dramatically and cannot be used in practice. Several researchers try to reduce the replicated packets along the network by using historic node meeting information [2, 5-7]. ProPHET [7] is using past node meeting information to compute the probability of meeting a node again. Nodes that are encountered frequently have higher probability to meet again and older contacts degrade over time. Messages are only replicated when the probability exceeds a threshold.

SimBet [8] is similar to our approach in that it also uses neighbor information to construct local graph to calculate graph properties of the encounter events. The authors use the concept of similarity and betweenness to develop SimBetUtil. In SimBet, when two nodes meet, they compute the SimBetUtil scores, and the packet is forwarded to the node with higher score. Here, the social networking feature helps to identify nodes that are more capable in meeting other nodes for message delivery. A potential drawback could occur when stable social relationships (so stable encounter patterns) do not occur in a particular DTN application. That is to say, the analysis of similarity and betweenness is sensitive to the mobility of a network. In SimBet, when the node with the highest score may have the highest probability to contact every nodes, so to deliver the message eventually. It is also possible that it could be hard for a message to be delivered outside a node with high SimBet score. For example, if there exists a node with high SimBet score near the source node and several lower SimBet score nodes close to the destination, it's difficult for the message to be delivered. If replying on the node with higher score to deliver message, SimBet will also require large buffer size.

B. Related Graph Theory: DTN can be treated as an undirected graph in which nodes represent the mobile stations and edges represent the communications between mobile stations. In this paper, we make use of the cut vertex concept from graph theory in order to find the nodes utilized to carry and deliver the messages. A cut vertex is defined as: a vertex in the graph such that removal of the vertex disconnects a connected graph. The concept of cut vertex can be applied to both directed and undirected graphs.

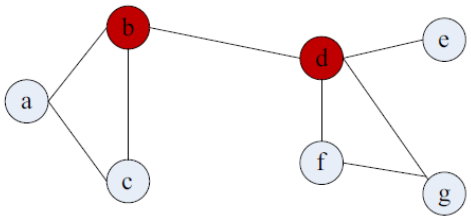


Figure 1. An undirected graph with two cut vertices b and d.

The nodes b and d in Fig. 1 are the cut vertices of an undirected graph. The removal of either node b or d will disconnect the graph. However, knowledge of the whole graph is required to compute cut vertices. In DTN, we cannot rely on the knowledge of the global network. We modify this concept and propose cut nodes which are cut vertices of a local sub-graph.

III THE PROTOCOL

A. Cut Node Computation: As we mentioned earlier, we cannot assume global knowledge of a DTN. Instead, every node in the network maintains the routing information to the nodes which are its neighbors.

Our definition of a cut node: when two nodes meet each other, they will exchange a list of their directly connected neighbors along with any known connections among those neighbors. Thus, they can construct a connected sub-graph G' (V', E') where V' is the vertex set which includes the two nodes and all nodes with direct connection to them and E' is an undirected edge set which represents the known connections among those nodes. We then compute the cut vertices of G' by using depth first search (DFS). We call a cut vertex found in G' a cut node of G. For example, node b and node h will construct a local graph G, depicted in Fig. 2, when they meet. From Fig. 2 we can see nodes b and h are cut vertices in that local graph and also become cut nodes according to our definition.

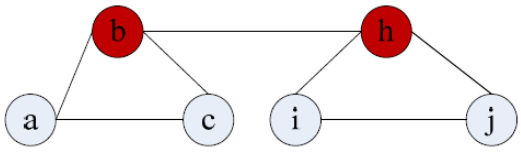


Figure 2. A local graph G' is constructed when node band h meet.

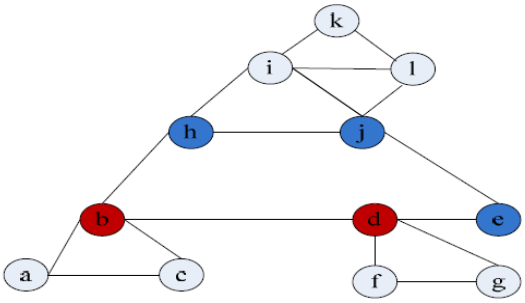


Figure 3. An undirected graph contains cut vertices b and d.

Fig. 3 shows some characteristics of cut nodes. Notice that nodes b and d, colored red are cut vertices of the whole graph and will be cut nodes, while nodes h, j, and e, colored blue, can also be cut nodes of the graph. The cut nodes can play important roles in delivering messages. For example, in Fig. 3, when nodes i, k, or l want to send a message to any of nodes a, c, f, or g, the message must go through one or more cut nodes. It is also important to note that although computing the cut vertices of a graph requires global knowledge of the graph, computing the cut nodes is performed with only a local sub-graph.

Here we also give a concrete example of the disadvantage of the SimBet algorithm where messages can be trapped at the node with the highest betweenness. Suppose in Fig. 3, node g wants to send a message to node k, it will forward the message to d first since node d has the highest betweenness and the similarity score is 0. Once node d has the message, there is no chance that it can further forward the message since both its neighbors b and e have lower SimBet scores. The message will stay at node d forever. However if we use cut nodes to forward the message, then it will be forwarded either to node h or j, and thus the message will move closer to its destination.

B. The Protocol: Every node in our protocol has a message vector, neighbor vector and cut vector. The definitions of these vectors are described as follow:

Message Vector: This vector stores messages which need to be forwarded. When the vector overflows, the oldest message will be replaced by newest message.

Neighbor Vector: This vector stores directly connected neighbors along with their neighbors.

Cut Vector: This vector stores all the cut nodes the node has met.

In our protocol, when a node wants to send a message to a destination node, it will check if the destination node is in its neighbor vector, if so it will try to deliver the message to the destination node directly. Otherwise, it will store a copy of the message in its message vector and forward the message to all of its directly connected cut nodes. Those cut nodes will then attempt to deliver the message in the same manner as the source node.

Algorithm 1 describes the communication between nodes m and n when they met each other. When node n receives a hello message from node m, it will look in its neighbor vector to determine if m is a new neighbor. If this is the case, it will also search the destinations of the messages in its message vector, and any message with destination m is delivered. At this point, node n also asks node m for its neighbor vector. Node m will send the neighbor vector back to node n.

Algorithm 1. Pseudo-code of node n

```
upon receive Hello message from node m do
    if newNeighbor(m) == true
        if messageVector.containDestinationOf (m) == true
            deliverMessages(m)
        requestNeighbor (m)

upon receive neighbor vector nv from node m do
    updateNeighbor()
    calculateCutVertex(bv)
    propogateCutNodes()
    if isCutNode(n) == true
     || isCutNodeNeighbor (n) ==true
      requestMessage (m)

upon receive message vector mv from node m do
    updateMessageVector (mv)
    confirmReceiving()
```

Node n will construct a local sub-graph after receiving the neighbor vector. Then it will compute the cut vertices of this local graph. If node n is a cut node, node m will forward all current messages it carries to node n. Node n then becomes a carrier of those messages and vice versa. If neither node is a cut node, then messages will be forwarded from n to any neighbor of m that is not a neighbor of n and is a cut node. If none of the cases exists, then both nodes will keep the message.

IV EXPERIMENTAL RESULTS

We have implemented our Cut Node Based Routing Protocol and also the Epidemic flooding protocol using ONE (Opportunistic Network Emulator) simulator. ONE simulator allows users to create scenarios based upon different synthetic movement models and real-world traces and offers a framework for implementing routing and application protocols.

The focus of the simulator is on modeling the behavior of store carry-forward networking, and hence we deliberately refrain from detailed modeling of the lower layer mechanisms such as signal attenuation and congestion of the physical medium. Instead, the radio link is abstracted to a communication range and bit-rate. These are statically configured and typically assumed to remain constant over the simulation. However, the context awareness and dynamic link configuration mechanisms can be used to adjust both range and bitrate depending on the surroundings and the distance between peers.

The performance metric chosen for comparison is Delivery Probability. Delivery probability is defined as the ratio between the number of messages successfully delivered to the destination and the total number of messages generated. Its value ranges between 0 and 1. If all the messages are successfully delivered to their corresponding destinations, then the delivery probability is 1. So, delivery probability is an important metric to compare the routing protocols for Delay Tolerant Networks.

A. Message delivery statistics: Fig. 4 shows the comparison between Epidemic and Cut Node based Routing protocols in terms of message delivery probability. The Epidemic Routing Protocol suffers from buffer overflow problem when the buffer size is less. But, our Cut node based Routing algorithm delivers more number of messages than Epidemic Flooding. The message delivery probability increases with increase in buffer size.

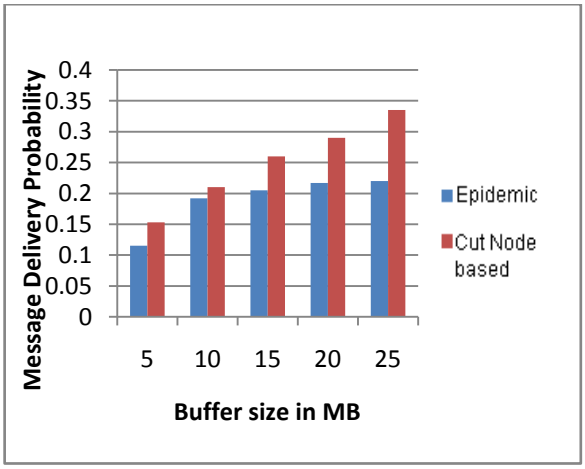


Figure 4. Message delivery statistics

B. Message Overhead Ratio: This metric is used to estimate the extra number of packets needed by the routing protocol for actual delivery of the data packets. It is defined as (Number of Packets Relayed - Number of Packets Delivered) / (Number of Packets Delivered). Epidemic routing creates replicas of messages for each time it forwards messages to the connected neighbors.

So, Epidemic routing puts lot of overhead on the network by the creation of large number of message copies. But, our cut-node based algorithm forwards messages only to the cut nodes. So, the overhead ratio is very less compared to Epidemic algorithm. The statistics are shown in the Fig. 5. We can also observe that the overhead ratio increases with the increase in buffer size as it prevents some of the messages to be dropped due to insufficient buffer space and hence those messages will increase the overhead ratio of the network.

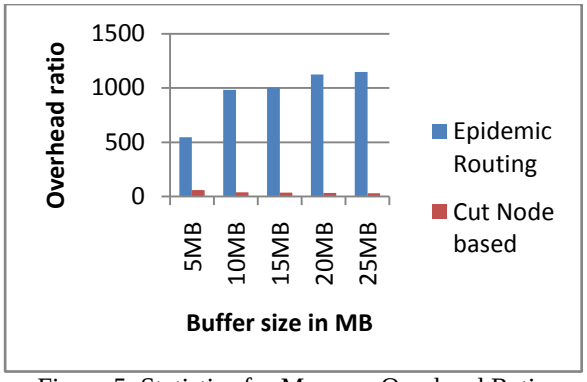


Figure 5. Statistics for Message Overhead Ratio

V CONCLUSION

In this paper, we propose an algorithm which utilizes graph theory to solve the routing problem in DTN. In our algorithm, when two nodes meet, they will exchange their neighbor information and construct a local sub-graph to compute Cut Vertices. We name those Cut Vertices as Cut Nodes. Those nodes have high probability to deliver messages outside the local cluster and hence also have high probability to forward messages to their destination. We simulate our algorithm using ONE simulator and compare our algorithm to Epidemic flooding. Our simulation results show that Cut Node Based Routing has high delivery rate.

VI REFERENCES

[1] J. Partan, J. kurose, and B. N. Levine, “A survey of practical issues in underwater networks,” in Proc. ACM WUWNet, 2006, pp. 17-24.
[2] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine “MaxProp: Routing for Vehicle-Based Disruption- Tolerant Networks,” in Proc. IEEE Infocom, 2006, pp. 1-11.
[3] J. Ott, D. Kutscher, “A Disconnection-Tolerant Transport for Drive-thru Internet Environments,” in Proc. IEEE INFOCOM, Miami, 2005, pp. 1849-1862.
[4] A. Vahdat, D. Becker “Epidemic routing for partially-connected ad hoc networks,” in Proc. Conference Name, Conference Location, 2000, pp.
[5] J. A. Davis, A. H. Fagg, and B. N. Levine “Wearable computers as packet transport mechanisms in highly-partitioned ad-hoc networks,” in Proc. IEEE ISWC, 2001, pp. 141-148.
[6] B. Burns, O. Brock, and B. N. Levine, “MV routing and capacity building in disruption tolerant networks,” in Proc. IEEE Infocom, 2005, pp. 398-408.
[7] A. Lindgren, A. Doria, and O. Schel´en, “Probabilistic routing in intermittently connected networks,” Lecture Notes in Computer Science, 2004, vol. 3126, pp. 239-254.
[8] E. Daly, M. Haahr, “Social network analysis for routing in disconnected delay-tolerant MANETs,” in Proc. MobiHoc 07, Montreal, Quebec, Canada, 2007, pp.32-4

Developing an IP core for PCI Using VHDL and its Verification

Dr. Rekha K. R, Ashwin Balakrishna
SJB Institute of Technology, Bangalore, Karnataka, 560060, India (VTU)

ABSTRACT **--- PCI is a computer bus for attaching hardware devices in a computer, these hardware devices are usually referred as the peripheral devices. PCI architecture was started by Intel in 1990, since then PCI has undergone a lot of changes and today PCI stands as one of the most versatile device with well established protocols to communicate with the peripheral devices. PCI allows the peripheral devices to directly access the system memory but it uses the bridge to connect to the front side bus and therefore to the CPU. The communication with the peripheral devices is established through cycles of bus transactions. The PCI bus transactions involve address phase followed by the data phase. The data phase may be from initiator to target or vice versa depending on the type of transactions taking place. A PCI interface will connect any generic devices to the PCI bus. The interface will create the communication between master and the slave devices. The interface will receive the command from the, master and send required response to the master and the slave as per the availability or status of both the master and the slave device. Generally most commands are provided by the master, our interface design includes read and write operations and will be able to communicate to register and RAM through the PCI.**

Index Terms **— PCI Protocol, PCI interface, VHDL.**

I. INTRODUCTION

PCI stands for peripheral component interconnect, as the name suggests PCI is a device which is mainly used to act as an interface between the master and the slave device. Use of devices such as PCI lead to standardization of the interface devices allowing these standard interfaces to connect to the a PC, capable of sustaining the high data transfer rates needed peripheral devices such as modern graphic controllers, storage media, network interface cards and other devices.[1]It may seem that PCI is a set of bus lines and computer users may have mistaken that PCI is just a set of electrical wires which help in transfer of data but PCI is actually a complete set of specification defining how different parts of the computer should interact. PCI bus is an improved bus for PC compatible computers. It has proved to be faster than previously introduces buses like the ISA, VESA. It has faster transfer rates and it can be expandable to both 32 bit and 64 bit. PCI covers most issues relating to computer interface. The PCI has distinct interface pins with specific functionality. We have to interface other devices as per the pin configuration of the PCI. The PCI bus has 4 main characteristics:

Synchronous: PCI bus operation uses one clock, which usually operates at 33 MHz or may choose to operate lower to save power.PCI can also operate at 66MHz if the hardware supports it.

Transaction/burst oriented: bus transactions are followed by an address phase followed by one or more data phases depending on the type of transaction taking place. The transaction is usually address phase followed by one or more data phases. Bus mastering: the operation of PCI works as a master slave relationship where usually the processor acts as an initiator.

Plug and Play: Host CPU/ host OS has the capability to identify the PCI devices connected to the board of the computer, hence it configures itself at the boot up. This feature is most important and is the reason for PCI devices popularity. PCI timing-PCI specifies timing related to its clock with a 33 MHZ clock, we have: 7ns0ns Tsu/Th (setup/hold) constraint on inputs 11ns Tco (click to output) on outputs. [2]

II. OBJECTIVE

The project mainly aims to provide communication between a master and slave device. Systems which include complex components always require interfaces which will act as mediators between to sub systems and allow them to communicate without any glitches. PCI is one such interfacing device which has become popular in today's systems. PCI will act as a mediator between the processor (master) and the peripheral device (slave). Both the master and slave will communicate only with the mediating interface device the PCI, but the versatile characteristics of the PCI device ensures that both the devices communicate successfully with each other.

III. INTERFACE I/O SIGNALS

Frame: It is the activation signal. When it is low the device indicates that it is ready to perform. [3]

Command or Byte Enable: It is a 4 bit signal, in the address cycle it indicates command and in the data cycle it indicates byte enable. [3]

Address: It is a 32 bit bus. It provides both address and data at the address phase and data phase respectively.

Initiator ready: It indicates that the master device is ready for transaction.

Target ready: Target ready is low when the slave device is ready to make transaction.
Device select: The target asserts device select (low) as acknowledgement to indicate that the address has been positively decoded.

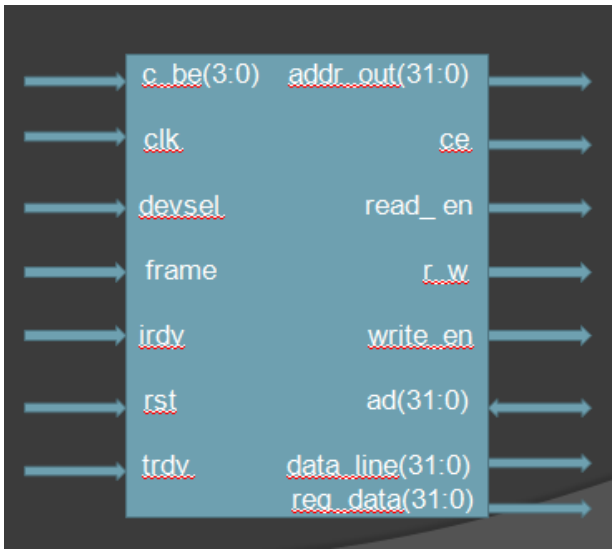


Fig.1. Interace block

IV. ADDRESS MAPPING

In our design, the method we have used to select the slave device is called the address mapping. PCI device has 32 bit address bus which means it can have 2^32 address which is 4294967296 number of addresses. Considering the address lines which are present in the PCI, it is evident that PCI has the ability to address many devices. Although PCI is capable of communicating with many devices at once due to electrical issues most of the PCI based systems usually restrict their communication to few devices.

V. SIMULATION ENVIORNMENT

In our system we consider a processor as the master of the system and a PCI target device as the interface. The master i.e. the processor provides the commands and indicates what type of transaction has to take place; basically the processor will act as an initiator. The transaction initiated by the processor will be received by the PCI target and further processing will be done as per the master's request.

VI. TIMING DIAGRAM

A. READ TRANSACTION:

The following timing diagram illustrates a read transaction on the PCI bus:

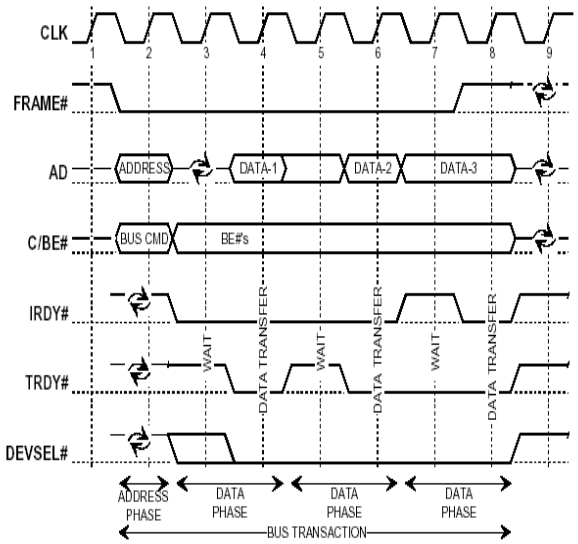


Fig.2. Read cycle of PCI.

Cycle 1- The bus is idle.
Cycle 2- The PCI initiates every transaction with an address phase to indicate which device it wishes to communicate with. This cycle is called the address phase where address of the device is received by the PCI bus from the processor. In this phase the AD lines indicate the address and the C/BE lines indicate the command or the type of transaction required to take place.
Cycle 3- The initiator tri-states the address in preparation for the target driving read data. Here the initiator will drive the valid byte enable signal C/BE indicating which are the valid data among the 32 bit data being transferred. IRDY# is low here indicating that the initiator is ready to make the transfer. Target will also assert the DEVSEL# low as an acknowledgment signal to indicate that the address has been correctly decoded.
Cycle 4- valid data is present at the AD lines. TRDY# is also low here indicating that the target is ready for the transfer of data. When the IRDY# and the TRDY# both are low the data transaction takes place here.
Cycle 5-Target will deassert TRDY#, making it high indicating it needs more time to prepare the next data transfer.
Cycle 6- The second data phase occurs when both the IRDY# and the TRDY# are low and again the data transaction takes place.
Cycle 7-The target provides valid data for the third data phase, initiator delays by indicating it's not ready for the transfer of the data.
Cycle 8- Initiator reasserts the IRDY# signal to indicate it is ready for the transfer of the data. Initiator captures the data provided by the target. Initiator drives the FRAME# high indicating the loast data phase.
Cycle 9-FRAME#, AD, and C/BE# are tri-stated, as IRDY#, TRDY#, and DEVSEL# are driven inactive high for one cycle prior to being tri-stated. [5]

B. WRITE TRANSACTION:

The following timing diagram illustrates a write transaction on the PCI bus:
Cycle 1- The bus id IDLE.
Cycle 2- The initiator asserts a valid address And places a write command on the C/BE# signals. This is the address phase.

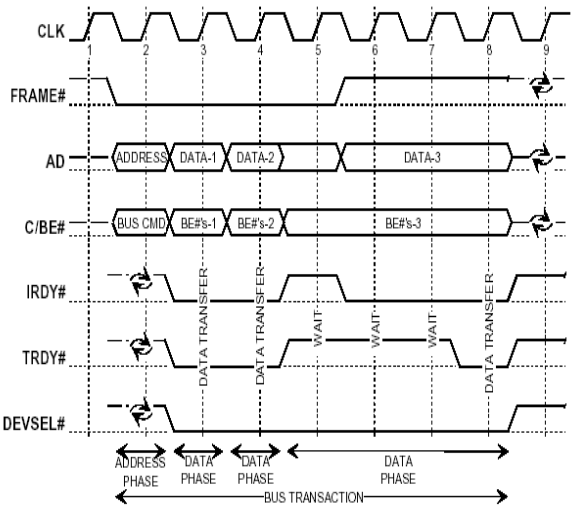


Fig.3. Write cycle of PCI.

Cycle 3- the initiator drives valid data and byte enable signals. The initiator asserts IRDY# low indicating valid write data is available. The target asserts DEVSEL# low as an acknowledgment it has positively decoded the address. The target drives TRDY# low indicating it is ready to capture data. The first data phase occurs the write data.
Cycle 4- The initiator provides new data and byte enables. The second data phase occurs as both IRDY# and TRDY# are low. The target captures the write data.
Cycle 5- The initiator deasserts IRDY# indicating it is not ready to provide the next data. The target deasserts TRDY# indicating it is not ready to capture the next data.
Cycle 6- The initiator provides the next valid data and asserts IRDY# low. The initiator drives FRAME# high indicating this is the final data phase. The target is still not ready and keeps TRDY# high.
Cycle 7- The target is still not ready and keeps TRDY# high.
Cycle 8- The target becomes ready and asserts TRDY# low. The third data phase occurs as both IRDY# and TRDY# are low. The target captures the write data.
Cycle 9- FRAME#, AD, and C/BE# are tri-stated, as IRDY#, TRDY#, and DEVSEL# are driven inactive high for one cycle prior to being tri-stated.[5]

VI. FINITE STATE MACHINES

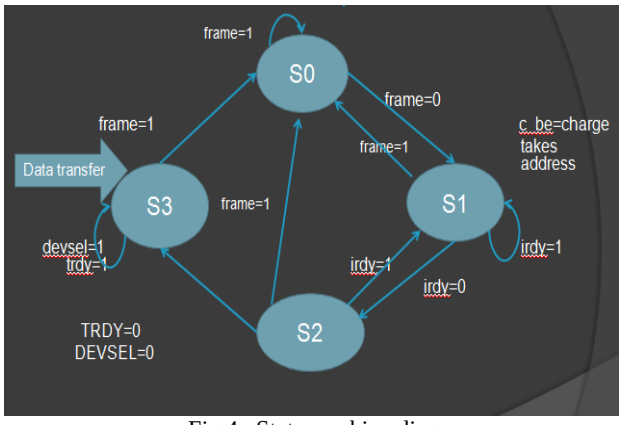


Fig.4. State machine diagram

S0=STATE_IDLE
S1=STATE_TAKE_ADDRESS
S2=STATE_TAKE_DATA
S3=STATE_LAST_DATA

A. STATE_IDLE:

In the idle state all the signals are reset to their initial states and if the system is completely idle here. PCI will wait for the response of the master for any operation to take place. PCI becomes active and ready for operation only when it detects that the FRAME # is low indicating the start of any operation be it configuration cycles or read and write cycles.

B. STATE_TAKE_ADDRESS:

The second state id the take address state where the address of the device to which master wishes to communicate is received by the PCI bus. Here the IRDY# is still at high state once the initiator is ready for the operation to take place the IRDY# is made low and this results in the change of state from the STATE_TAKE_ADDRESS state to the next state.

C. STATE_TAKE_DATA:

The third state is the take data state and if it is found that the DEVSEL# and the TRDY# are low it indicates that the device address has been decoded correctly and further data transaction takes place.

D. STATE_LAST_DATA:

The last and the final state of the data transaction is the STATE_LAST_DATA state where the last data is transferred and this is indicated by the FRAME# signal when it goes high. Hence this completes one full cycle of the data transaction and the PCI reverts back to the IDLE state.

VII. SIMULATION AND IMPLEMENTATION

For simulation process, we have used Xilinx 10.1. For synthesis and verification we have used Xilinx 10.1 and ModelSim 6.4 SE.The simulation results captured by the ModelSim software are shown below:

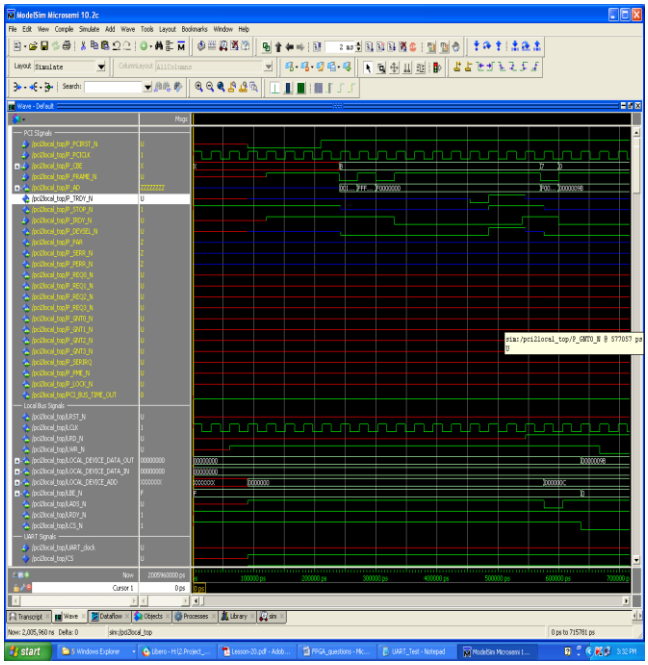


Fig.5. Waveform – Configuration cycle.

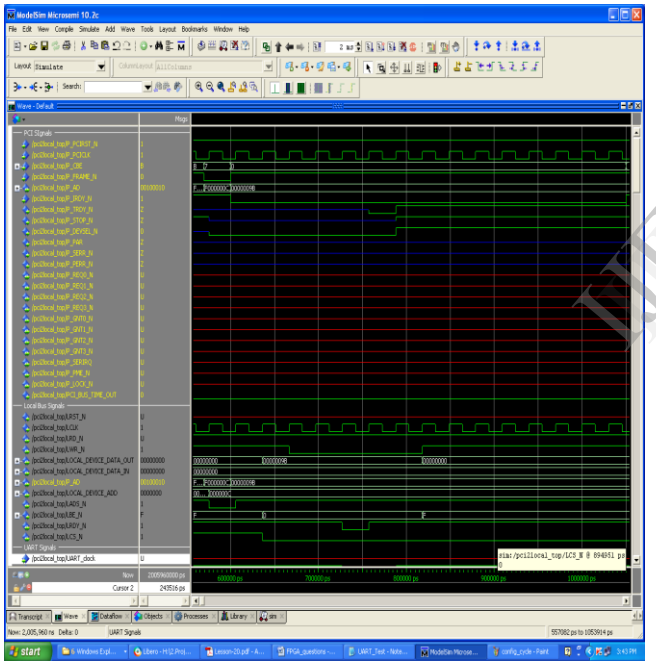


Fig.6. Waveform -Data transaction 1.

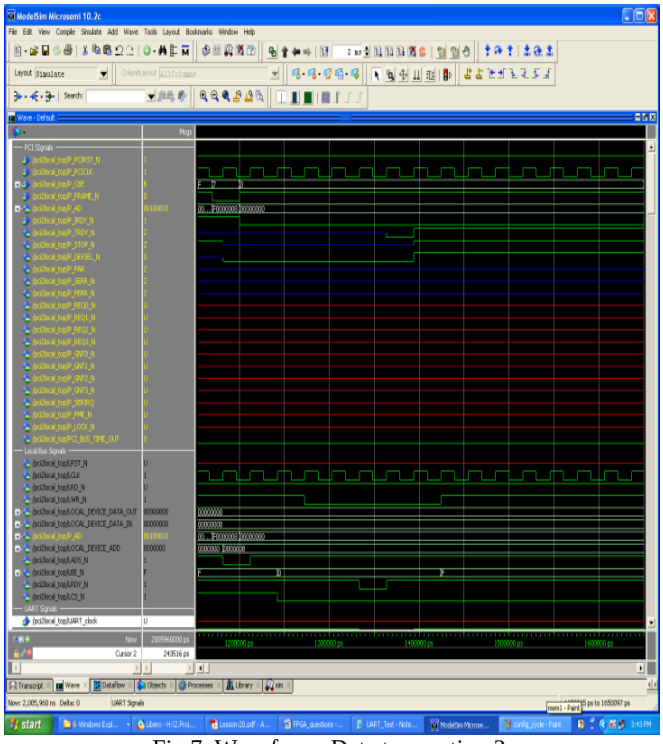


Fig.7. Waveform -Data transaction 2.

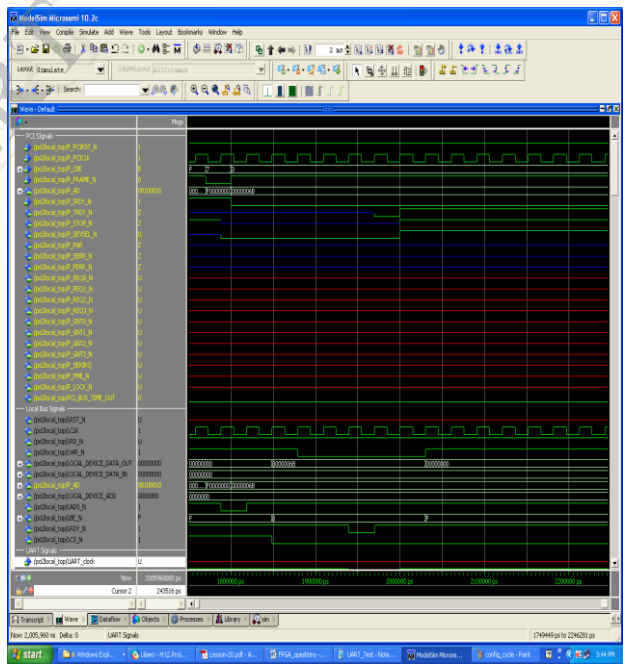


Fig.8. Waveform -Data transaction 3.

VIII. CONCLUSION

Today's computer systems, with their emphasis on high resolution graphics, full motion video, high bandwidth networking, and so on, go far beyond the capabilities of the architecture that ushered in the age of the personal computer. Today's PC s demand high transfer rate and deal with sophisticated peripheral devices hence interfaces like PCI is very much necessary. It is a well thought out standard with a number of forward looking feature that should keep it

relevant well into the next century. PCI was originally conceived as a mechanism for interacting with the master and a slave in a system and to serve as an interface in the system. But today it stands as a well defined device which provides glitch free communication between the master and the slave device.

Though in this project we have not illustrated the communication of the PCI device with any peripheral device the simulation results proves that successful communication will be possible using the PCI. We hope that this project will help in further PCI related invention. PCI has been evolving and adapting since its birth and has proven one of the best interface and today we also have many advanced version of the PCI to match higher speeds and be compatible with more sophisticated systems. Nonetheless PCI has retained its popularity and has played a remarkable role with its significant speed and reliability.

ACKNOWLEDGMENT

The authors would like to acknowledge HOD and faculty of SJBIT (ECE Department) for their consultation and advice throughout the research work.

REFERENCES

[1] Tech-pro.net. (n.d.). Retrieved August 5, 2010, from tech-pro.net: http://www.tech-pro.net/intro_pci.html

[2] Fpga4fun. (2010, September 07). Retrieved September 07, 2010, from fpga4fun: http://www.fpga4fun.com/PCI1.html

[3] PCI Local Bus Technical Summary. (2010, October 10). Retrieved October 10, 2010, from techfest: www.techfest.com/hardware/bus/pci.htm

[4] Figure Block Diagram of Static RAM Table Truth Table. (2010, July 27). Retrieved July 27, 2010, from docstoc:http://www.docstoc.com/docs/4219029/Figure-Block-Diagram-of-Static- RAM-Table-Truth-Table

[5] PCI Bus Timing Diagram. (2010, August 17). Retrieved august 17, 2010, fromsilverhawk:http://silverhawk.net/notes/tutorials/hardware/pcitiming.html

[6] Altera DE1 board. (n.d.). Retrieved December 10, 2010, from terasic:http://www.terasic.com.tw/cgibipage/archive.pl?Language=English No=83

Development of Log Files Analysing Tool for Bmc Remedy Ar System

Lavanya G S[1], and Prof Pushpa H G[2]
[1]M.tech Student, Department Of Computer Science, SBMJCE, Bangalore, India
[2]Head of the Department, Department Of Computer Science, SBMJCE, Bangalore, India

***Abstract*—Log files contain large amount of data which is used while debugging an application. Analyzing these software log files helps during testing and troubleshooting. A log file analyzer for BMC Remedy Action Request System (AR System) is presented in this paper. Log files grow very quickly depending on user activity and type of logging. Many times these log files grow to 200-300 Mbytes within 4-5 minutes. To physically go through such huge log files is very painful and consumes lot of time as well. To resolve this, application that can parse through these log files and trace the execution of workflows for given user and present the same in a graphical & intuitive interface is presented.**

***Key words*—Log files,AR System, workflow, graphical interface**

I. INTRODUCTION

The logging activity is mainly been used as a debugging aid which is particularly vital to distributed applications. Large software systems often keep log files of events. Such log files can be analyzed to check whether a run of a program reveals faults in the system. The runtime information that is being recorded in the log files provides an easy and very powerful debugging approach. Pausing program execution at a breakpoint disturbs the time order of interacting with subsystems as the software runs in a controlled environment [1]. Hence this logging activity sometimes turns out to be the most feasible debugging method. In automated log file analysis data from the log files is extracted. This requires decoding the log file syntax and interpreting data semantics. The output of this phase is used by an organization for further processing. The log data extractors can be developed using programming languages targeting log file formats [2]. A generic scheme for interpreting elements of a log file is desired instead of repeating this process for each log file format and using suitable data structure for further processing. Almost every tool present in the market generates some form of a log which can be used to identify software defects that can be hardly detected in conventional testing.

AR System is a professional development environment that leverages the best practices of the IT Infrastructure Library (ITIL) and provides a foundation for Business Service Management (BSM) solutions. In addition, log file analysis also helps in extracting vital information from AR servers and in security monitoring. The output of the AR System log analyzer is a tree filled with the information of interest for the particular case. The log files generated in AR server are analyzed for many purposes. In particular, the goal is to verify the functional conformance of the application with a given specification for each workflow in the AR System.The log file entries are observed to confirm that the application generates desired outputs at intended instances. When an application malfunctions in production, the only trace available for developers is to investigate the cause, which more often, leads to the examination of the application log file.

Powerful log files analysis tools are required with capabilities to monitor the execution of different workflows, the order of execution, number of users connected to server, timestamp related to each workflow object etc.for AR System. Correlating data extracted from a monitoring tool with an application log can reveal valuable information caused by important application events. Despite the benefits of it, log file analysis is a labor intensive and error prone activity when performed manually. Furthermore, it generally demands domain and tool expertise which is an expensive resource. Therefore, automating at least a part of the process has significant importance. The first step in automated log analysis is automatic extraction ofrelevant information from log files. Doing this in a generic way is a challenging task given that different log files generated in the AR server have different structures and formats.

II. RELATED WORK

Log files have great importance in AR System since they help us in identifying the root-cause of an incident. If thereis a production release, logs might be full of errors and these errors must be ordered according to their impact. These errors are then solved one by one till a clean log file is obtained. AR LogAnalyzer is an existing program that is used to analyze AR System SQL and API logs only. The output provides a breakdown of execution times for each API call and SQL statement and, for API logs, a breakdown of idle times by thread. Collectively, this output is used to pinpoint performance trouble spots and to help guide AR System Administrators generally in their efforts to do performance tuning.

This tool requires a command-line script with no built-in GUI interface or prompting. It does not come with specific guarantees concerning the operation of the utility or the accuracy of its output. AR LogAnalyzer does not provide any graphical representation for the given log files; neither does it provide color codes for easy debugging. As this tool is

operational from the command prompt it is not user friendly. To overcome these draw back, log analyzer tool is developed which provides reliable results along with tree structure representation of the contents of the log files.

III. SYSTEM ARCHITECTURE

The main components of an AR System application include:

- Forms - The main AR System application component that users interact with is a form. Each form is composed of fields.
- Menu—Menus are lists that you create to guide the user in entering information in fields on forms.
- Workflow—Forms provide the mechanism to structure the data captured and menus offer options for specific field data, additional components—active links , filters, and escalations—act on the data to automate business processes, or workflow. These components trigger actions in response to execution options that is defined.
    - An active link is an action or group of actions performed on the client. Active links are triggered by user actions in a form.
    - A filter is an action or group of actions performed on the AR System server. Filters are used to enforce business rules and to ensure system and data integrity.
    - An escalation is an action or group of actions performed on the server at specified times or time intervals.

The main components of the parser include user interface, AR server, log parser and finally the graphical view of the logs processed.

AR Server: A log file records everything that goes in and out of a particular server. The information is frequently recorded chronologically, and is located in the root directory, or occasionally in a secondary folder, depending on how it is set up with the server. The only person who has regular access to the log files of a server is the server administrator. A log file is generally password protected, so that the server administrator has a record of everyone and everything that wants to look at the log files for a specific server. Since these servers are built with proprietary technologies, only way to debug any application issues is to go through the log files generated and find out which workflow fired/not fired. Problem with this approach is that these log files grow very quickly depending on user activity and type of logging. Also, since these are server side logs, they contain workflows executed for all users and not just the user who is having issues. To physically go through such huge log files is very painful and consumes lot of time as well. In the process productivity of developer also gets affected.

Log Parser: Log file parser is needed that can parse through these log files and trace the execution of workflows for given user and present the same in a graphical & intuitive interface. This would greatly reduce time & effort spent by developers in troubleshooting application issues.

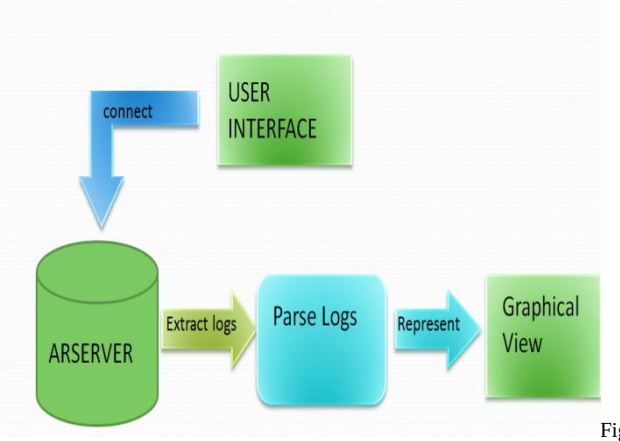


Figure 1: System Architecture

User Interface: User Interface module allows user to connect any AR Server located in any geographical area of an organization. User should know the AR Server name to login into server. Once user logs using valid credentials into server, he is able to give particular log file name to parse and to get all information related to the log file. User request the system through User Interface. User request provides the transaction id in particular field.

Extract and parse: In this module program extracts particular log file mentioned by user and give it to parsing. During parsing, log file will be read line by line and separates active links and filters present into two different text files. Based on the one of the conditions, i.e., transaction id (TID) or user, parser will execute and provide a graphical tree representation of the same. Here we are using pattern matching as our key concept to separate active links and filters and also to find which workflows triggered and which did not.

Graphical View: The execution of workflows is presented in a tree structure which helps in easy debugging of the application. User is Capable of viewing workflow based on particular user or TID assigned to a user. The tree structure also indicates whether selected workflow is executed or not, whether selected workflow is enabled or disabled and the actions performed if this workflow is executed.

IV. IMPLEMENTATION

The function of workflow is to process the data captured in forms in accordance with the business needs. In AR System, workflow automates a company's processes through the use of active links, filters, and escalations. In general, workflow can be defined as the set of processes that a company uses to run itself. Each of these actions will be recorded in the log files.

The application developed can parse through these log files and trace the execution of workflows for given user and present the same in a graphical & intuitive interface. This would greatly reduce time & effort spent by developers in

troubleshooting application issues. The application developed has the capabilities:

- Capable of parsing proprietary log files generated by BMC Remedy Applications
- Capable of parsing log files up to 1 GB of size.
- Capable of identifying workflow execution per user.
- Display execution of workflows in a graphical & intuitive manner.
- It has the ability to highlight workflows on User Interface(UI) and shows following details about the workflow when highlighted:
    - Whether selected workflow is executed or not.
    - Whether selected workflow is enabled or disabled.
    - Workflow execution condition.
    - Value of parameters set while executing this condition.
    - Actions performed if this workflow is executed.

The proposed system working architecture is as shown in figure 2. Here the user extracts the logs from BMC AR server by connecting to it, the extracted log files are then parsed and represented in a form which is useful to the developer in troubleshooting the issues.

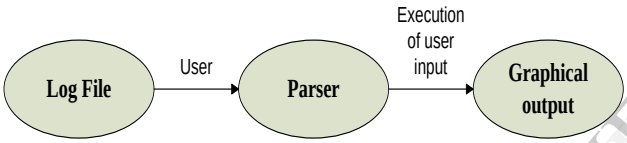


Figure 2: Proposed System

The central servers in an organization will be having huge amount of log files in terms of GB or TB those get generated every day. Our program will get access to AR Server only when particular server is accessed by authenticated user. The program will search for particular log in AR server database. If found it retrieves file name and load into the parser for parsing a log file. If it is not found then display error message to the user on UI.

The Parser is responsible for extracting the log file from AR server database into parsing tool. It separates remedy objects; active links, filters and other workflows. To separate the remedy objects, pattern matching using regular expressions approach is used. The log files are also processed using different condition based on workflow names that have triggered, users logged on to the server and also the transaction id used by particular users. The tool also provides information regarding the number of users currently connected to the server and their last modification time. For a given workflow present in the log file the parser returns a detailed list of actions performed, the workflow objects primary form name, the execution order of the workflow etc.

Graphical View: is responsible for viewing all workflows according to the user triggering the actions. User can view the workflows related to only active links or filters. Users can also view the workflow of different users to trace the execution of each workflow present in the server. This view is designed like tree structure with an expandable option which when clicked reveals the actions performed if any. Each active link/filter name is colored green or red based on the passed or failed qualification. All the disabled workflows are represented in blue. Each specific action is represented by unique color in the tree structure. Thus, by just observing the graphical representation developers can easily identify the cause of error based on the qualifications used to trigger each workflow. Totally this gives complete representation of what user has done from his system and the course of actions that has followed during the transactions.

V. RESULTS

The main page of the tool is as shown in figure 3. The user provides the log file generated by BMC AR Server. The log file is selected by clicking on browse button through a file chooser. Once the file is selected user selects the criteria to parse the file as shown in figure 3.

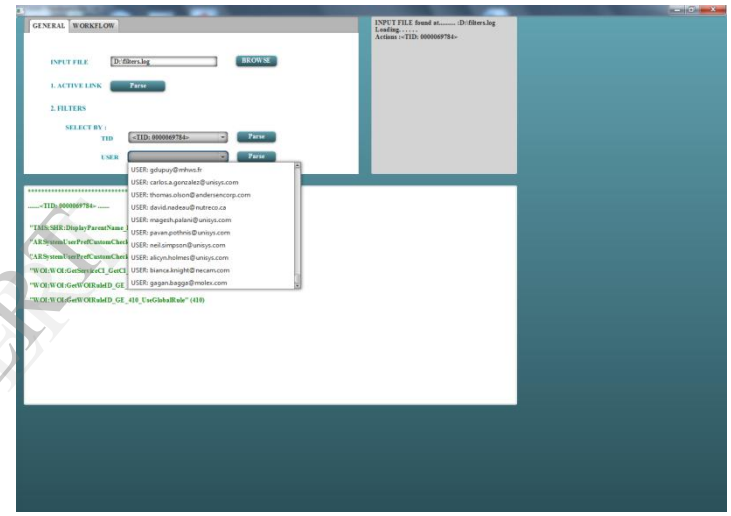


Figure 3: Selection of criteria to parse the log file

The parsing of log file is done based on Thread ID and also based on Users present in the log file. As seen in Figure 4, parsed graphical representation of log file based on Thread ID is done.

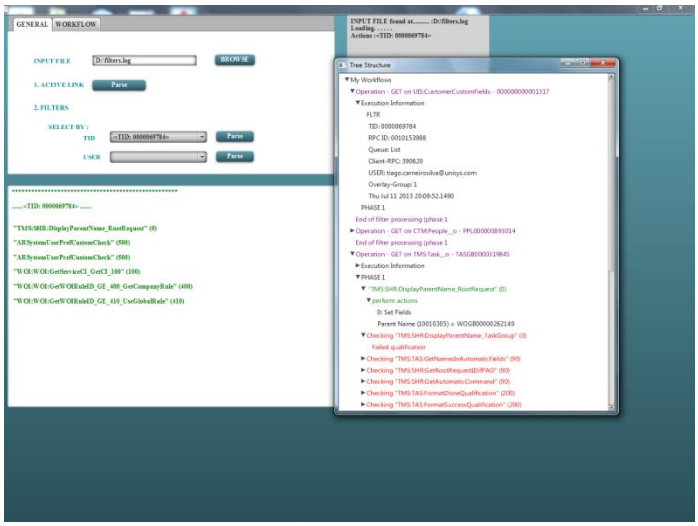


Figure 4: Generation of tree structure based on the TID

By selecting the user text box, the workflow for that particular user can be represented graphically in similar fashion as shown in figure 4. For each selection the names of the forms that contains the workflow names is displayed in the text box area. A console area is also provided where the background action performed is listed as shown in figure 5.

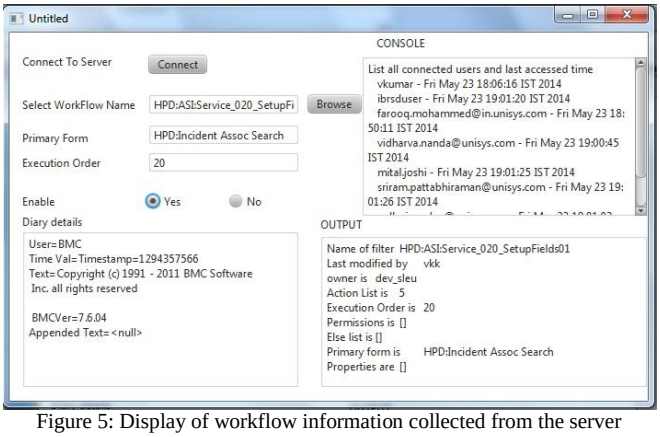


Figure 5: Display of workflow information collected from the server

VI. CONCLUSION

This tool will be useful in corporate environment for folks working with remedy log files for resolving the problems related to application with time constraint involved. The time spent on manually going through huge, confusing log files is saved. This time could be fruitfully spent on the solution of the problem rather than finding the cause of the problem. This indeed reduces the effort of the developers in solving the issues and increases the productivity in an organization.

REFRENCES

[1] Dapeng Liu; Shaochun Xu; Huafu Liu "Using Log Files as Knowledge Bases "Industrial and Information Systems (ICIIS), in 6th IEEE International Conference onAdvanced Applied Informatics, 2011 On page(s): 130 – 135
[2] P.W.D.C. Jayathilake, "A Novel Mind Map Based Approach for Log Data Extraction" in 6th International Conference on Industrial and Information Systems,2011, ICIIS 2011, Aug. 16-19, 2011, Sri Lanka
[3] B. Beizer, Software Testing Techniques, 2nd Edition, 2 Sub. IntlThomson Computer Pr (T), 1990
[4] "Log4j." [Online]. Available: http://logging.apache.org/log4j/1.2
[5] J. H. Andrews, "Testing using log file analysis: tools, methods, and issues," in 13th IEEE International Conference on Automated SoftwareEngineering, 1998. Proceedings, 1998, pp. 157–166.
[6] J. H. Andrews, "Testing using log file analysis: tools, methods and issues," Proc. 13th IEEE International Conference on AutomatedSoftware Engineering, Oct. 1998, pp. 157-166.
[7] James H. Andrews, "Testing using Log File Analysis:Tools, Methods, and Issues"
[8] J.H.Andrews. Theory and practice of log file analysis. Technical Report 524, Department of Computer Science,Universityof Western Ontario, May 1998
[9] BMC Remedy Action Request System 7.6.04 Concepts Guide
[10] L.K. Dillon, G. Kutty, L. E.Moser, P.M.Melliar-Smith, and Y. S. Ramakrishna. Graphical interval logic for specifying concurrent systems. ACM Transactions on Software Engineeringand Methodology, 3(2):131–165, April 1994.

Dynamic Resource Allocation Using Load Balancing

Sanchaya S
Department of Computer Science and Engineering,
Jain Global Campus, Jain University,
Jakkasandra Post, Kanakpura Taluk,
Ramanagar District-562112

Madhu B.R
Assistant Professor
Department of Computer Science and Engineering,
Jain Global Campus, Jain University,
Jakkasandra Post, Kanakapura Taluk,
Ramanagara District-562112

Abstract: **In cloud computing environment, resource allocation under bursty workloads can be achieved by enhancing the load balancer. Load balancer is a method to distribute workloads across multiple computers, central processing units, disk drives or many other resources. Resource utilization and maximizing the throughput can be achieved by minimizing the response time and avoid overloading. In the existing load balancing system the algorithm does not consider burstiness as well as current resource utilization in user demands. Hence the existing algorithm which is static in nature can be thought of made dynamic. This improves the system performance and provides faster response time.**

Keywords: Cloud computing, Load Balancing, Burstiness.

1. INTRODUCTION

Cloud computing is an on demand service since it offers dynamic flexible resource allocation for reliable and guaranteed services in pay as you use manner to the public. Multiple cloud users can request number of cloud services simultaneously when necessary. So there must be a provision that all resources are made available based upon the request made by the user in efficient manner to satisfy their needs.

In cloud platforms allocation of resources takes place at two stages. The first stage involves the application to be uploaded to the cloud, the load balancer starts assigning the requested instances to the physical computers in order to balance the load across many physical computers. The second stage involves the application to receive multiple incoming requests, here each requests should be specifically assigned to the application instance in order to balance the computational load across many instances of the same application. Amazon Elastic Compute Cloud uses elastic load balancing to control and handle the incoming requests.

The presence of burstiness in the user workloads usually causes the degradation of the application performance. In order to satisfy the peak user demands, load balancer usually does not consider the case of bursty arrivals and hence results in performance degradation. Burstiness also causes load unbalancing in clouds and as a results degrades the system overall performance.

As a result, finding out the burstiness and providing high quality of service along with system availability is important and challenging as well. When the resources are over-utilized, it results in increased response time. Similarly when the resources are under-utilized, it results in wastage of resources.

Load Balancing is necessary for efficient operation in distributed environment. Cloud computing is a well known platform for providing storage of data in an inexpensive manner that is available over the internet and hence load balancing has necessarily become one of the important and interesting topics in the research fields.

When the number of requests is being generated simultaneously, balancing the load is necessary for achieving better user satisfaction as well as to utilize the resources based on the availability. There many algorithms those are available for providing efficient mechanism and to enhance the cloud performance. Hence provides satisfying and efficient services to the user.

2. RELATED WORK

Burstiness has been known as an important characteristic of traffic in communication networks and has fueled much research over the past two decades. Recently the presence of burstiness has also been identified in a variety of settings, including enterprise systems grid storage systems and file systems. The impact of burstiness has been examined and reported in [5].

Tai Jianzhe et.al [5] implemented a new smart load balancer by adjusting the tradeoffs between randomness and greediness in the site selection process.

Shreyas Mulay., et al [10] implemented the cluster sorting of servers for load balancing. Hence with help of cluster sorting technique requests are handled easily handled by server clusters.

Rashmi K. S., et al [2] A load balancing algorithm has been proposed to avoid deadlocks among the Virtual Machines (VMs) while processing the requests received from the users by VM migration.

Ram Prasad P., et al [3] implemented the idea regarding "Load balancing in cloud computing system" which includes distributed servers along with high fault tolerance, availability scalability and so on.

Mishra, Ratan et.al [4] An ant colony optimization has been proposed to initiate the service load distribution under

cloud computing architecture. The pheromone update mechanism has been proved as a efficient and effective tool to balance the load. This modification supports to minimize the make span of the cloud computing based services and portability of servicing the request also has been converged using the ant colony optimization technique. This technique does not consider the fault tolerance issues.

3. PROPOSED SYSTEM

The proposed load balancing system is a combination of ARA online algorithm and Enhanced Equally distributed algorithm. This proposed load balancing system has an advantage over the previous algorithms in terms of its response time, efficiency and throughput. The design of the proposed work is shown in Figure1.

In this method, the total available servers are initially grouped into a set of 3 servers each known as clusters (Data Center). This is because, when the servers are grouped into clusters, the sorting of clusters will be easier and quicker compared to sequential server sorting. Secondly, clusters acts as backups for each other i.e., if one of the cluster is over loaded, the request will be handled by other clusters until that cluster gets back to normal state. So, this increases the efficiency and response time of the system.

Logic of operation–This system involves two stage sorting i.e. one at the cluster level and other at the server level. Each one of them is associated with a variable called Cluster counter variable (CCV) and Server counter variable (SCV) respectively. These variables will be updated automatically as the cluster and server status changes. Thus, load balancer will sort the cluster and server in descending order of their values.

Cluster counter variable defines the maximum number of request that the cluster can handle. E.g. If CCV is 300, then the 3 servers in that cluster can handle 300 requests simultaneously (loads may or may not be equally distributed within the cluster).Similarly, Server counter variable defines the number of requests that each servers can handle. E.g. If SCV is 100; it means that the server can handle 100 requests simultaneously.

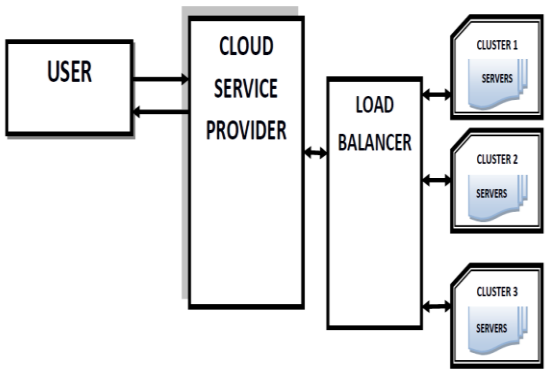


Figure1: Load Balancing in Cloud Computing

Method of operation – Initially, when the user requests arrive from the client, the load balancer (LB) counts the number of incoming requests. Later the LB sends a query to the clusters to know its status. Once the CCV is received, LB arranges the cluster in the descending order of their CCV values. That means, the cluster with maximum request handling capability will be at the Priority 1 level and the cluster with least request handling capability will be at the priority K level, where K refers to the total no. of clusters in the LB system.

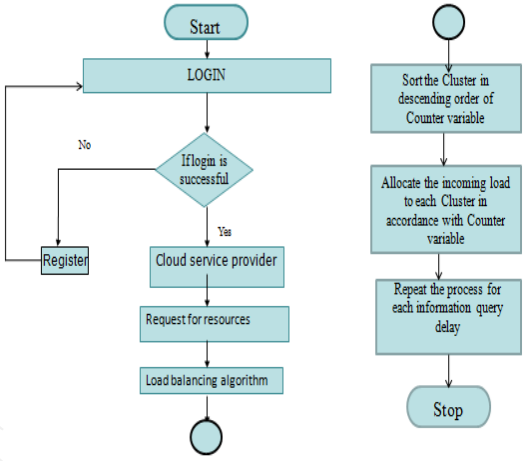


Figure 2: Proposed Flow Chart

The Figure 2 shows the flow of the project. During this time, the LB also receives the server counter variable, from each cluster and sorts the servers in descending order of their SCV value. That means, inside each cluster, the server with maximum request handling capability will be placed on the top and one with least request handling capability will be placed at the bottom. Later the prediction algorithm is executed to know the type of incoming requests. If the request is a bursty type, then the LB will automatically switch to Random mode that is the requests will be randomly allocated to the sorted servers without observing the kind of requests. This improves the response time but the performance output may not be up to the expectation.

If the incoming request is weak bursty or idle type, then LB switches to greedy mode, that is, servers will be allocated which best suits the requests. This improves the performance output and there will be no delay in response as the requests are few in number.

Depending on the no. of incoming requests, the no. of clusters will be varied. E.g. suppose each cluster can handle 500 requests simultaneously and there are 10 clusters in a cloud under one LB, if the request size is 4000, then it could be handled by 8 clusters. So, the remaining 2 clusters will be kept in passive mode, so that any other LBs in the cloud can utilize that clusters for their services. This increases the efficiency of the server utilization and also the throughput in cloud computing. Once again if the request increases, then the clusters return to the parent LB and provide service for them.

4. EXPERIMENTAL ANALYSIS AND RESULTS

The registered client requests for the resources available in the cloud. The requests are redirected to the servers based on the availability. The cluster with more number of servers available for providing services is given the highest priority.

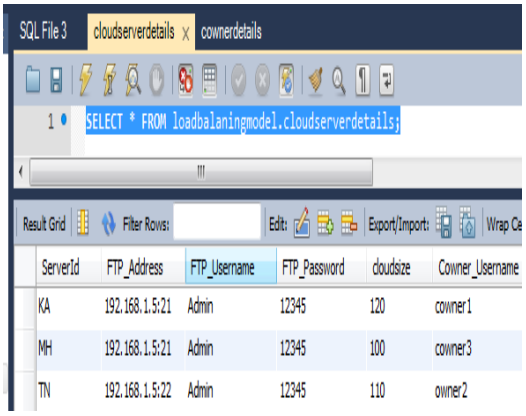


Figure 3: Servers KA, MH, TN are available with cloud size 120 MB, 100MB, 110MB respectively.

In Figure 3, the server details along with the cloud size is shown. Since KA is the server with highest priority based upon the cloud size, when a request is made by the client the server KA is used for allocation.

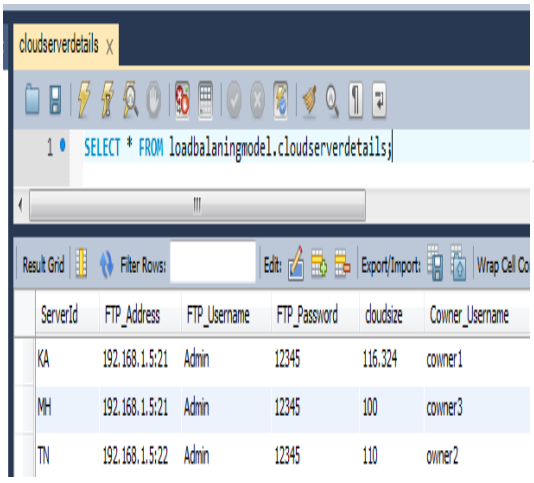


Figure 4: Updated server details after resource allocation

In Figure 4, the size of the cloud is being reduced after the allocation of the resource is made by the server with the highest priority (i.e. KA).

5. CONCLUSION AND FUTURE WORK

In this paper the description is about allocating the resources under bursty workloads. The proposed work focuses on allocation of resources to the servers based on server sorting. The resources are efficiently utilized by assigning the counter variables based upon the availability of the server. Hence efficiency of the system can be achieved. The future work involves the concept of cluster sorting based on the availability of servers.

REFERENCES

[1] Naimesh D. Naik and Ashilkumar R. Patel "Load Balancing Under Bursty Environment For Cloud Computing."International Journal Engineering Research and Technology (IJERT) ISSN: 2278-0181 Vol. 2 Issue 6, June – 2013.

[2] Rashmi K. S,Suma V.,Vaidehi M., "Enhanced Load Balancing Approach to Avoid Deadlocks in Cloud" in Special Issue of International Journal of Computer Applications (0975 – 8887) on Advanced Computing and Communication Technologies June 2012.

[3] Padhy Ram Prasad, & P. Gautam Prasad Rao. "Load Balancing in Cloud Computing System" at Department of Computer Science and Engineering National Institute of Technology, Rourkela-769 008, Orissa, India May, 2011.

[4] Mishra Ratan and JaiswalAnant. "ANT Colony Optimization :A solution of load balancing in cloud in International Journal of Web & Semantic Technology IJWET Vol.3, No.2, April 2012.

[5] Tai Jianzhe, Zhang Juemin, Li Jun, Meleis Waleedand MiNingfang "ArA: Adaptive resource allocation for cloud computing environments under bursty workloads".In the 30th IEEE International Performance Computing and Communications Conference.

[6] Shreyas Mulay and Sanjay Jain "Enhanced equally distrusted load balancing for cloud computing Volume 2,Issue no 6,June 2013.

[7]http://www.computer.org/csdl/proceedings/pccc/2011/0010/00/06108060-abs.html

[8]http://www.ijert.org/view.php?id=1621&title=architecture-for-distributing-load dynamically-in-cloud-using-server-performance-analysis-under-bursty-workloads

[9]http://www.collaborative.com/uploads/Drive%20OnDemand%20Performance%20Testing%20with%20Cloud%20Computing%20and%20Proactively%20Meet%20Your%20Market%20Needs.pdf

[10] http://warse.org/pdfs/ijmcis01112012.pdf

Evaluation of Polyphase Filter Architecture for pulse detection and measurement

Jeevitha T
Digital electronics and communication,Dayananda Sagar College of engineering, Bangalore.

Abstract: This paper addresses the design of polyphase filter architecture for pulse detection and measurement as applied to the current EW(Electronic Warfare) systems. The filter is designed for a rectangular input of length 256.The comparison is made based on the simulation results.

I. INTRODUCTION:

An EW system is used to protect military resources from enemy threats. The field of EW is recognized as having three components. They are:

1.Electronic Support Measure: collects the information on the electronic environment.

2.Electronic Countermeasures: jams or disturbs the enemy systems.

3.Electronic Counter countermeasures: protects the equipment against ECM.

Wideband and narrowband digital receivers are the two important types of systems used in EW and communications. However Wideband receivers have a very wide instantaneous input bandwidth of about 1GHz or larger. Another difference between the two is that wideband intercept receivers operate in an environment where the information of the input signal is unknown. Another major difference between the two is EW receivers output pulse descriptor words(PDWs), which describe the characteristics of the detected signal.

The basic block diagram of the digital wideband EW receiver is as shown figure:

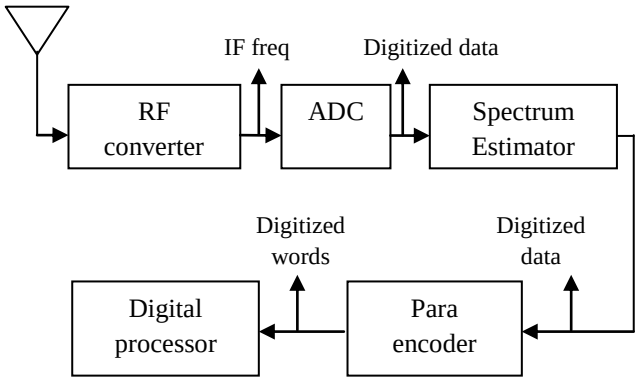


figure1 :digital wideband EW receiver

The antenna at the left captures pulsed RF signals that are generated by most radars. RF ranges from 2GHz-100GHz, but for radar , the most popular frequency range is from 2GHz-18GHz. Next an RF converter down converts the frequency signal into a lower intermediate frequency(IF) signal so that the EW receiver can more easily process the same bandwidth signal at a lower frequency. The signal then proliferates into the spectrum estimator. Spectrum estimation is commonly done using Fast Fourier Transform. The digital representation of the frequency spectrum is then output to the parameter encoder which generates PDWs.

The performance of an EW receiver mainly depends upon its update rate or its time resolution. Improved time resolution results in quicker and accurate TOA and pulse width measurements. For a digital EW receiver , the spectrum estimator is the limiting factor for the update rate.

The proposed method in this research for increasing time resolution is to use decimation in frequency domain. By decimating in the frequency domain, the computational load on the spectrum estimator is greatly reduced. This decimation in the frequency domain is implemented through channelized filtering structure.

Polyphase filtering is a multirate signal processing operation leads to an efficient filtering structure for hardware implementation. Polyphase filtering parallelizes the filtering operation there by reducing the computational load on the spectrum estimator.

The paper is organized as follows, first a brief introduction on the topic is given. Next the design procedure is discussed, followed by the algorithm and then based on the simulation results conclusions are drawn.

II. DESIGN PROCEDURE:

The basic block diagram of the Polyphase FFT filter architecture is as shown in the figure:

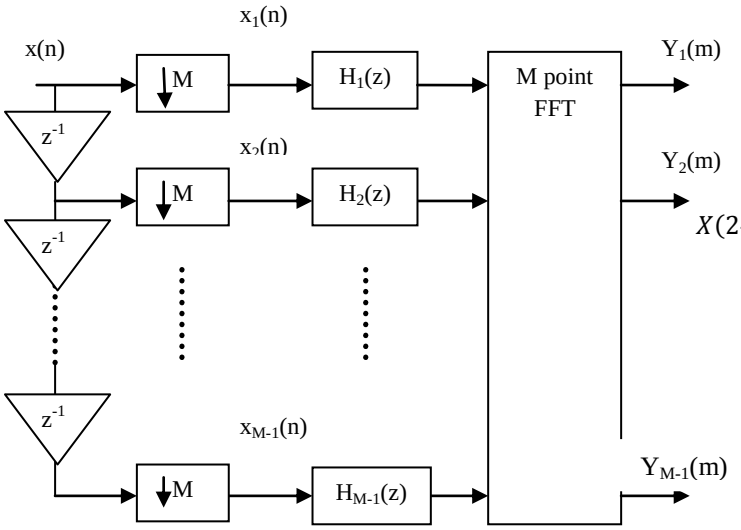


figure 2 : polyphase FFT architecture

The incoming N data samples are distributed into M branches and an M point FFT is performed. The number of points M is selected based on the time resolution required. The decimation results in gaps in the frequency domain. Hence each FFT filter must be widened to cover the gaps in the frequency domain. This is done by applying time domain window to the incoming data.

III. Algorithm: Decimation in Frequency domain through polyphase filtering:

The frequency domain decimation operation is best viewed visually in the context of a filter bank. A filter bank is simply the filter response of each sub band in the spectrum overlapped on each other. Let us assume that one frequency component from the FFT output is equivalent to one filter output at a specific point in time from a filter bank. Let us assume that the 256 –point FFT is decimated by a factor of 8.

A 256 point FFT can be written as:

$$X(k) = \sum_{n=0}^{N-1} x(n).e^{\frac{-j2\pi nk}{N}} \quad (1)$$

where N=256.There are 256 outputs in the frequency domain. If every 8th output is kept and the other outputs are discarded , the resulting outputs are k=0,8,16,……,248. There are a total of 32(256/8) outputs. These outputs can be written as:

$$X(0) = \sum_{n=0}^{255} x(n)$$

$$X(8) = \sum_{n=0}^{255} x(n).e^{\frac{-j2\pi 8n}{256}}$$

$$X(16) = \sum_{n=0}^{255} x(n).e^{\frac{-j2\pi 16n}{256}}$$

$$\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet$$

$$X(248) = \sum_{n=0}^{255} x(n).e^{\frac{-j2\pi 248n}{256}} \quad (2)$$

Let us arbitrarily choose a frequency component k=16 and rewrite it by decomposing it into its polyphase components. The result is:

$$X(16) = \sum_{n=0}^{255} x(n).e^{\frac{-j2\pi 16n}{256}} = \sum_{n=0}^{255} x(n).e^{\frac{-j2\pi 2n}{32}}$$

$$= [x(0) + x(32) + x(64) + \ldots.. + x(224)]$$

$$+[x(1)+x(33)+x(65)+\ldots\ldots..+x(225)].e^{\frac{-j2\pi 2}{32}} +$$

$$[x(2)+x(34)+x(66)+\ldots\ldots\ldots+x(226)].e^{\frac{-j2\pi 2*2}{32}} +$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$+[x(31)+x(63)+x(95)+\ldots\ldots\ldots+x(225)].e^{\frac{-j2\pi 2*31}{32}} \quad (3)$$

In the above equation the relation $e^{-j2\pi n} = 1$ where n is an integer. Now let us define a new quantity y(n) as:

$$y(n) = x(n)+x(n+32)+x(n+64)+ \ldots\ldots\ldots\ldots + x(n+224)$$

$$= \sum_{n=0}^{7} x(n + 32m). \quad (4)$$

where n=0 to 31. This y(n) represents the values in the bracket of the above equation. Each y(n) contains a total of 8 data points.

Using these y(n) values , the FFT results from (3)can be rewritten as :

$$X(0) = \sum_{n=0}^{31} y(n)$$

$$X(8) = \sum_{n=0}^{31} y(n).e^{\frac{-j2\pi n}{32}}$$

$$X(16) = \sum_{n=0}^{32} y(n).e^{\frac{-j2\pi 2n}{32}}$$

• • • • • • • • • •

$$X(248) = \sum_{n=0}^{31} y(n).e^{\frac{-j2\pi 31n}{32}} \quad (5)$$

All these equations can be written into one equation as:

$$X(8k) = \sum_{n=0}^{31} y(n).e^{\frac{-j2\pi kn}{32}} \quad (6)$$

where k=0,1,2,…..,31 and n=0,1,2,…..,31.

The output X(8k) can be relabeled as Y(k), thus the above equation can be written as:

$$Y(k) = \sum_{n=0}^{31} y(n).e^{\frac{-j2\pi kn}{32}} \quad (7)$$

This equation represents a 32 point FFT. In order to obtain the outputs of a 256 – point FFT decimated by 8, a 32 point FFT can achieve the goal. Thus the design of the FFT can be simplified. The input must be manipulated, however in order to obtain the desired result.

A general statement without further proof will be presented here. If one wants to perform an N point FFT and the outputs in the frequency domain are decimated by M, one can achieve the goal by performing an N/M – point FFT. A new input format y(n) must be built first. The generalization of the y(n) can be written as:

$$y(n) = \sum_{n=0}^{M-1} x(n + mN/M). \quad (8)$$

where n=0,1,2,……. (N/M)-1.The outputs in the frequency domain can be obtained as:

$$X(n) = \sum_{n=0}^{\left(\frac{N}{M}\right)-1} y(n).e^{\frac{-j2\pi kn}{N/M}} \quad (9)$$

Hence the design of the FFT can be simplified tremendously.

If the filter response is plotted for the (7), the resulting filter response will have very low side lobe level.The resulting filter bank has many gaps in the frequency domain.As a result each filter response must be widened.This is accomplished by applying a window function to the incoming data.

The window function that is used here is a Parks McClellan window function.

The input data x(n) will be modified by the window function h(n).Here, h(n) instead of w(n) is used for the window function because h(n) will be used to represent the impulse function of the filter. The resulting data $x_m(n)$ used as the input of the FFT can be written as:

$$x_m(n) = x(n).h(n) \quad (10)$$

where n= 0,1,2,…., 255.As stated previously the outputs are decimated by 8. Under this condition the modified data can be used in (4) to find the y(n) as:

$$y(n) = \sum_{n=0}^{7} x_m(n + 32m) = \sum_{n=0}^{7} x(n + 32m).h(n + 32m) \quad (11)$$

where n= 0,1,2,…..,31. A 32 point FFT is performed on these y(n) values , and the individual filter response as well as the filter bank is plotted to observe the improved filter shape.

IV. SIMULATION RESULTS:

The simulations are done by means of Matlab. A rectangular window was used as the input.As stated earlier the sample size was considered to be N= 256 and the decimation factor as M= 8. The sampling frequency was assumed to be 1350MHz.The filter designed to widen each filter output was generated using FDATOOL Matlab for the following parameters:

Fs= 1350MHz;

Fp= Fs*(M/N) = 42.1875 MHz;

Fs= 2*Fp =84.375MHz;

wp= 1/(N/M/2) = 1/64;

ws = 2*wp =1/32;

density factor =20;

N=256;

1. For the FFT of a rectangular window described by the equation 1, the individual filter response and the filter bank with overlapping and non-overlapping are as shown below respectively:

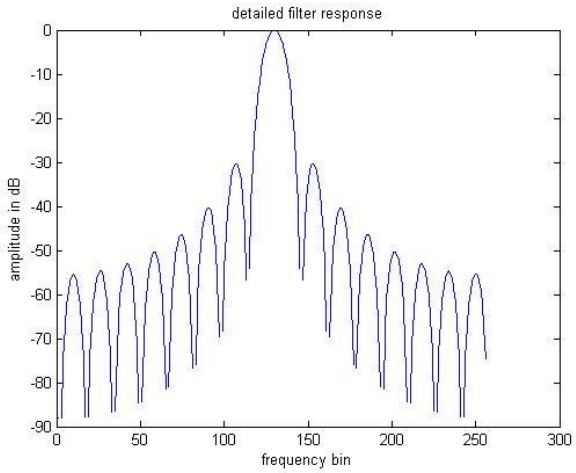


figure 3: 256 FFT individual filter response

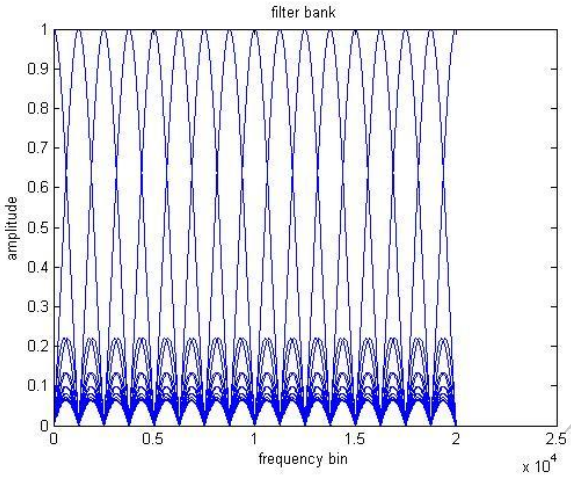


figure 4: 256 pt FFT overlapped filter bank

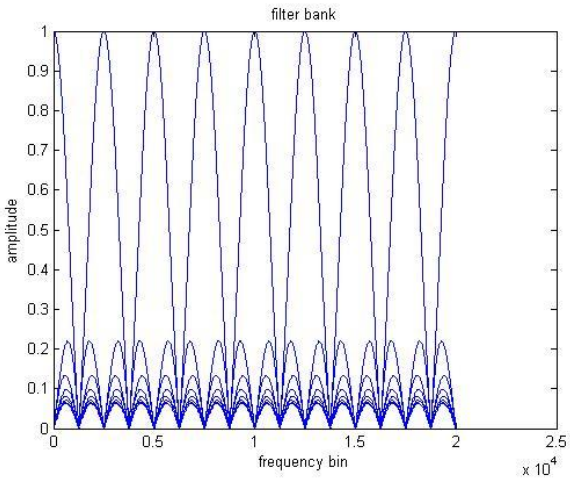


figure 5 : 256 pt, FFT non-overlapping filter bank.

It can be seen that the sidelobe levels are very high also the transition band is not sharp. As a result this type of a response is undesirable.

2. After applying the decimation in frequency domain algorithm to the same input and performing a 32 point FFT on the output values and plotting the individual filter response and its filter bank respectively:

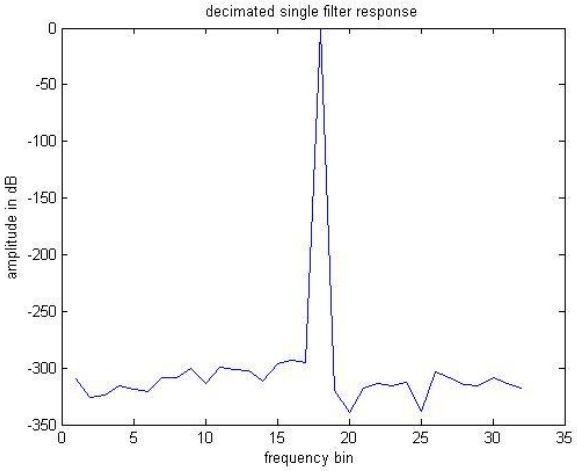


figure 6: 32 pt decimated FFT individual filter response.

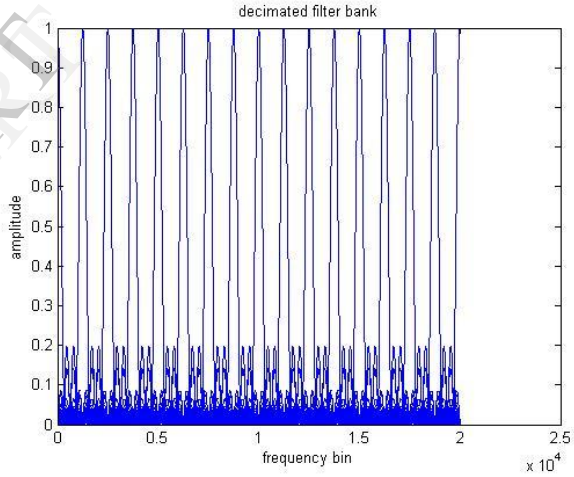


figure 7: 32 pt. decimated FFT filter bank

The filter bank as seen in the figure has many gaps resulting from the decimation operation. If an input signal frequency falls within one of the gaps, the receiver will not detect the signal. As a result this filter shape is not acceptable, each filter must be widened. This is accomplished by applying a window function to the incoming signal.

3. In order to widen each filter while suppressing the side lobes , a windowing function is applied to the incoming data.The magnitude and the impulse response of the filter that is used is as shown respectively below:

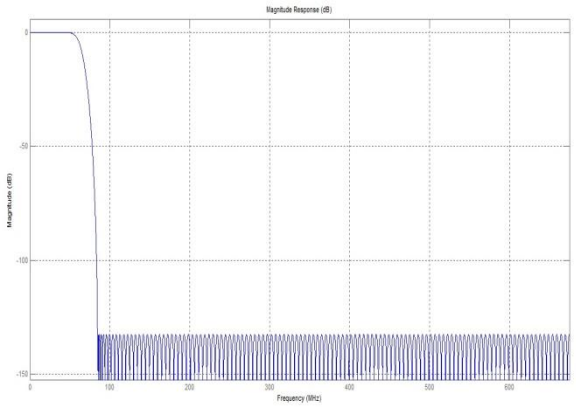


figure 7: magnitude response of the lowpass filter

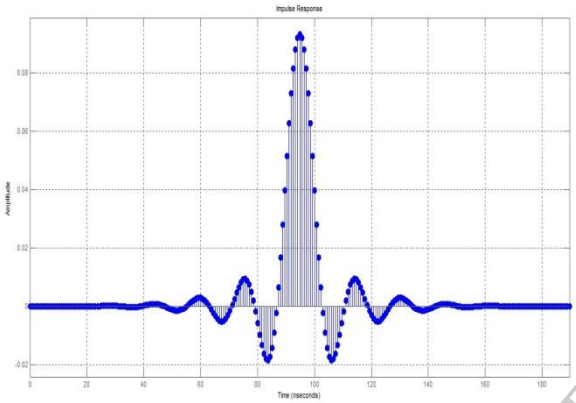


figure 8 : impulse response of the lowpass filter

If the 32 pt. FFT is performed on the resultant y(n) values , the FFT will modulate the prototype lowpass filter co-efficients and the input data to generate a filter bank of 16 unique filters.. The resulting individual filter response and the filter bank is as shown respectively:

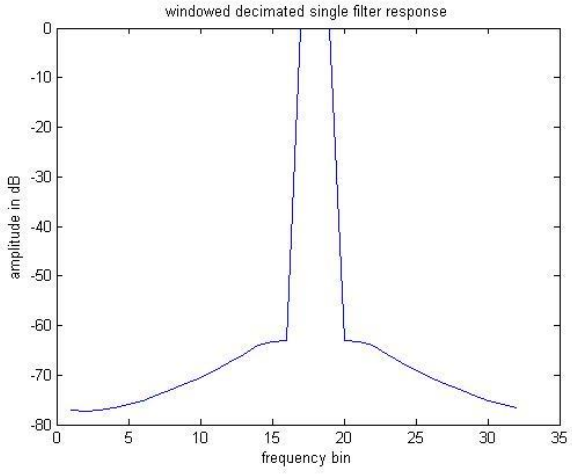


figure 9 : 32 pt. windowed decimated individual FFT filter response.

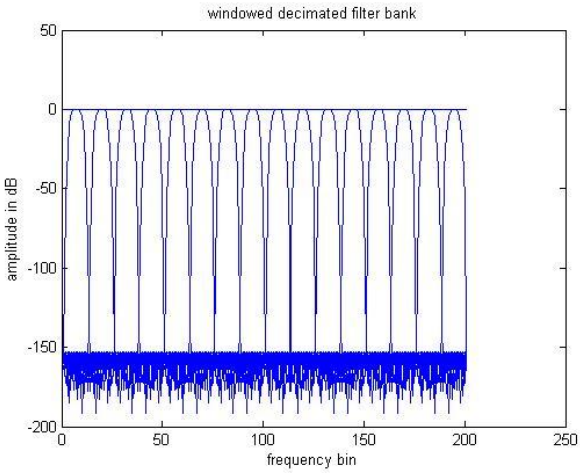


figure 10 : 32 pt. windowed decimated FFT filter bank.

V. Conclusion:

A significant advantage of the decimation in frequency domain is the improvement in the time resolution, which improves the TOA and pulse width measurements.

For the original 256 point FFT with the sampling frequency of 1350MHz the time resolution is 189ns. For the decimated case with 32 point FFT the time resolution is 23ns. There is as eight fold improvement in the time resolution.

For the windowed filter bank , there are no missed frequencies, but its ability to resolve two frequencies that are close together is limited. For this reason the decimation in frequency domain has a direct effect on frequency resolution.

VI. REFERENCES:

[1]. Abhijit S Kulkarni, Vijesh P , Hemant V Paranjape, Dr. K Maheshwara Reddy ,"Approaches towards the implementation of multibit digital receiver using FFT",Defence science journal, vol.63,No.2, march 2013, @ DESIDOC.

[2]. Buxa , Peter. Parameterizable Channelized wideband digital receiver for high update rate Wright state university, M.S Thesis 2003.

[3]. James B.Y.Tsui , Digital techniques for wideband receivers , 2nd edition. Norwood, MA:Artech House,2001

[4]. Richard lyons, Understanding digital signal processing , Pearson Education , 2004.

[5]. James B.Y.Tsui, Special design topics in digital wideband receivers , Norwood, MA: Artech House, 2010.

[6]. Ronald Crochiere, Lawrence Rabiner, Multirate digital signal processing.

Fast Human Detection using Histogram of Templates and Haar-like Features

Chethan k and ShahlaSohail
Dayanandasagar college of engineering, Bangalore,
Karnataka, 560078, India

***Abstract*—Pedestrian detection in images is still a problem with view and posture variation. In this paper, combination of a novel feature named histogram of templates (HOT) and Haar-like features are used as descriptors for feature extraction in an image. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. HOT features are extracted from every pixel of an image which are meeting various templates for a pre-defined formula. Extracted features from both the methods are provided to support vector machine (SVM) classifier for training and classification.**

***Index Terms* — Histogram of templates, Haar-like features, SVM.**

I. Introduction

Human detection in images is very important in the area of image based sensing, applications such as surveillance, pedestrian detection, robotics[1]-[6]. Many methods have been introduced for human detection in common views and simple background, it is still a problem in the situations like complex background, different views and postures.

The two main problems for designing any human detection system is feature representation and classifier design. The Human detection can be divided into techniques which require background subtraction or segmentation and techniques which can detect humans directly from the input without such pre-processing[7]. Background subtraction techniques usually find the foreground object from the image and then classify it into categories like human, animal, vehicle etc., based on shape, color, or motion or other features. Direct techniques operate on (features extracted from) image patches and classify them as human or non-human. We can also classify techniques based on the features which are used to classify a given input as human or not. These features include shape (in the form of contours or other descriptors), color (skin color detection), motion, or combinations of these.Some of the feature descriptors are Haar-like features[6], HOG[8], Local binary patterns(LBF)[9],HOG-LBF[10], Granularity-tunable Gradients partition(GGP) descriptors[11]. According to most recentresearches performance and accuracy in detection can be improved by combining different feature descriptors[12].

II. Previous Work

Paul Viola, Daniel Snow, Michael J Jones [6] describes pedestrians detection system that integrates image intensity information with motion information. Use detection style algorithm that scans a detector over two consecutive frames of a video sequence. The detector is trained using Adaboost algorithm to take the advantage of both intensity information and motion information. Intensity information is calculated by finding histogram of image and Motion information can be extracted from pairs or sequences of images by measuring the differences between region averages at various scales, orientations, and aspect ratios. Generalization of the Viola Jones features which operate on the differences between pairs of images in time is used. Using optimized image processing routines the time taken to detect human in an image can be greatly reduced.

Christoph H Lampert [15] introduced Efficient Sub-window Cascade (ESC), a divide and conquer for accelerating the evaluation of classifier cascades for object detection in natural images. The ESC algorithm starts with a single window in stage 1 that contains all possible object locations. Depending on the quality bounds, the window is either accepted as a whole, rejected as a whole, or split into disjoint parts that are separately processed further. Accepted windows are advanced to the next classifier stage or returned as detection if they already were in the last stage. By using an internal representation by set of regions instead of individual regions, ESC can discard large fractions of the potential candidate locations with few classifier evaluations. Thereby it reduces the computational effort compared to the standard way of cascade evaluation for object detection, in which one applies the classifier cascade exhaustively to every candidate region in the images.

Y. Mu, S. Yan, Y. Liu, T. Huang, and B. Zhou [9] introduces Local Binary Patterns (LBP) for human detection. Existing LBP descriptors does not suite for human detection, due to its high complexity and lack of semantics consistency. For this, Paper proposed two variants of LBP which are Semantic LBP and Fourier LBP. Among two popular approaches for human detection this paper uses sub-window based method for feature extraction from an image. Here each neighbor pixel is compared with the center pixel, once whose intensities exceed the center pixels are marked as 1 otherwise 0. From this get binary code and convert it to decimal form for further calculations.

Q. Ye, Jiao, and B. Zhang [16] uses HOG with multi-scale windows for feature extraction. Different sizes of square image blocks are used and slides over entire image to get features. The extracted features are fed into a cascade adaboost to train the classifier, here classifier used was two stage classifier. Drawback of this method is cannot detect pedestrian in crowed scenes.

Shaopeng Tang and Satoshi Goto [17] uses concept of every pixel of image various templates are defined, each of which contains the pixel itself and two of its neighboring pixels. If the texture and gradient values of the three pixels satisfy the pre-defined formula, the corresponding template for this formula.This extracted features are passed to the SVM classifier for classification of pedestrians. Results shows that Histogram of Templates(HOT) feature is more discriminative than HOG feature for the same training method.

III. Proposed Method

This method gives overview of our feature extraction chain, which is summarized in Fig.1. This section is divided into A. Feature extraction B. Classifier

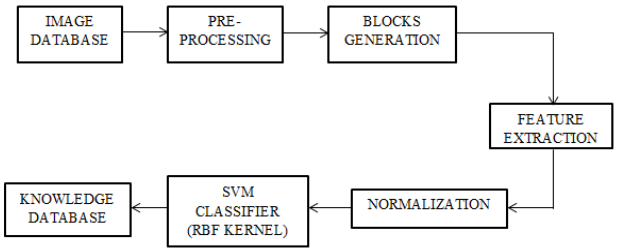


Fig .1. Flow chart for Human detection

A. *Feature Extraction*

In pattern recognition and in image processing, feature extraction is a special form of dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant then the input data will be transformed into a reduced representation set of features. Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

A Haar-like feature[6] considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. For example, let us say we have an image database with human faces. It is a common observation that among all faces the region of the eyes is darker than the region of the cheeks. Therefore a common haar feature for face detection is a set of two adjacent rectangles that lie above the eye and the cheek region. The position of these rectangles is defined relative to a detection window that acts like a bounding box to the target object.

In the detection phase a window of the target size(basic haar set as shown in Fig. 2) is moved over the input image, and for each subsection of the image the Haar-like feature is calculated. This difference is then compared to a learned threshold that separates non-objects from objects. Because such a Haar-like feature is only a weak learner or classifier (its detection quality is slightly better than random guessing) a large number of Haar-like features are necessary to describe an object with sufficient accuracy. In the Viola–Jones object detection framework, the Haar-like features are therefore organized in something called a classifier cascade to form a strong learner or classifier.

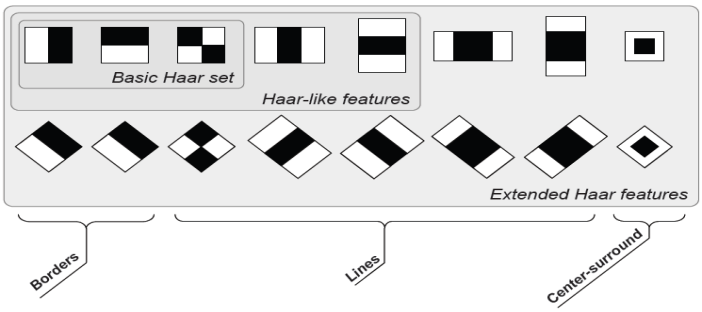


Fig .2. Basic Haar set for feature Extraction

For every pixel of an image, various templates are defined, each of which contains the pixel itself and two of its neighboring pixels [17]. If the texture and gradient values of the three pixels satisfy a predefined formula, the central pixel is regarded to meet the corresponding template for thisformula. Histograms of pixels meeting various templates are calculated for a set of formulas, and combined to be the feature for detection. Some templates are given to define the special relationship of three pixels in Fig. 3.

These templates are used in some formulas. The texture information and the gradient information are also used in these formulas, to give a concrete definition of this feature. The formulas are designed to capture the shape of the human body, and have reasonable computation complexity. For texture information, two formulas are given as following

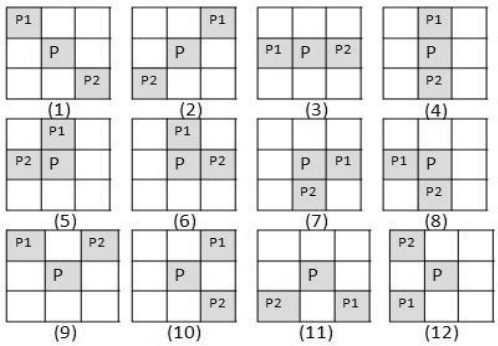


Fig. 3. Templates defining special relationship between three pixels.

$$I(P) > I(P1) \;\&\&\; I(P) > I(P2) \qquad (1)$$

For each template, if the intensity value of *P* is greater than the other two, it is regarded that the pixel *P* meets this template. It can capture the pixels that have the greatest value in one template, and the histogram of pixels that satisfy each template in a sub window can reflect the properties of local part of human body well. For each sub window, the number of pixels meeting each template is calculated to get a histogram as shown in Fig. 4. For example, eight templates are used to extract the feature

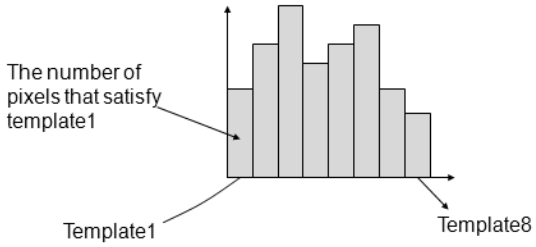


Fig 4. Example of histogram of template for one formula; 8 templates are used, and they correspond to 8 bins. The value of each bin is the number of pixels that meeting corresponding template.

$$k = = \arg_I \max\{ I(P_i) + I(P1_i) + I(P2_i)\} \qquad (2)$$

The sum of intensity values of three pixels in template *k* is greater than the values of other templates; it is can be regarded that *P* meets template *k* . A histogram can be calculated by using formula . By using this formula, we could find the template that has the greatest sum. They can be regarded as the basic unit of human body shape and the shape of human body can be represented well. For the gradient magnitude information, there exist similar formulas.

$$Mag(P) > Mag(P1) \;\&\&\; Mag(P) > Mag(P2) \qquad (3)$$

$$k = = \arg_i \max\{ mag(P_i) + mag(P1_i) + mag(P2_i)\} \qquad (4)$$

Eight templates are usually used to extract the feature, so for each formula, an eight-dimensional vector can be obtained. These vectors are combined together as the final feature as shown in Fig. 5.

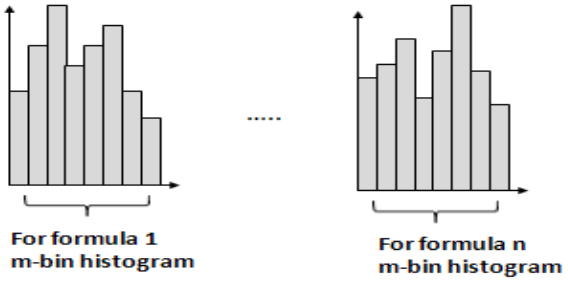


Fig. 5. Final HOT feature for a sub window

B. *Radial Bias Function Kernel SVM*

The inseparable data can be made separable by mapping original input space into a high-dimensional dot product space called the feature space [14], [18]. Mapping is done by a kernel function which is given by

Gaussian RBF kernel : $\Phi(r) = \exp(-r^2 / 2\sigma^2)$ for some $\sigma > 0$

Support vector machines(SVM) are used to find the particular hyper-plane that maximizes the margin of separation and finding such a separating hyper-plane which is optimal. The inner product kernel is a function that is used to find the optimal hyper-plane for SVM network.

Inner product kernel: $K(x, x_i) = \Phi^T(x)\,\Phi(x_i)$

Where Φ is a set of transformation functions, $x_i$is the i th training sample and x is the input sample. The input sample is mapped onto a feature space and searched for the optimal hyper-plane. The architecture of SVM is as shown in fig 6.

IV. Results

Database for training SVM is formed from selecting images from INIRIA dataset [1], [8] which consist of

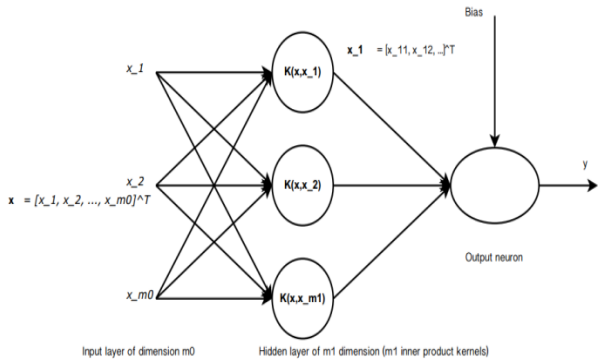


Fig. 6. SVM Architecture

2478 positives and 12180 negatives. For testing purpose images from the same dataset is considered. Figure 7 shows the human examples of two subsets

(a) (b)

Fig. 7. Human samples of two subsets from INRIA Dataset

By training SVM with samples of different subsets and using the knowledge database provided from SVM training for testing results in very good detection of humans in images. Figure 8 shows the detection of humans in images obtained from this experiment.

Fig. 8. Detection examples obtained from this experiment

V. Conclusion

Human detection in images using SVM as classifier and HOT and Harr-like features as descriptors is implemented using Matlab R2010a software. The proposed method can detect the human in images with various postures, clutter backgrounds, human shadows. By combining two feature descriptors the human detection accuracy in the images is improved.

References

[1]. Qixiang Ye, Zhenjun Han, Jianbin Jiao and JianZhuang Liu, "Human detection in Images via Piecewise Linear Support Vector Machines", IEEE Transactions on Image Processing, Vol. 22, No. 2, February 2013.

[2]. Y. Xu, D. Xu, S. Lin, T. X. Han, X. Cao, and X. Li, "Detection of sudden pedestrian crossings for driving assistance systems," *IEEE Trans.Circuits Syst. Video Technol.*, vol. 42, no. 3, pp. 729–739, Jun. 2008.

[3]. P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part based models," *IEEETrans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.

[4]. M. Enzweiler and D. M. Gavrila, "Monocular pedestrian detection: Survey and experiments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 12, pp. 2179–2195, Dec. 2009.

[5]. R. Xu, B. Zhang, Q. Ye, and J. Jiao, "Cascaded L1-norm minimization learning (CLML) classifier for human detection," in *Proc. IEEE Int.Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 89–96.

[6]. P. Viola, M. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance," *Int. J. Comput. Vis.*, vol. 63, no. 2, pp. 153–161, 2005.

[7]. Neeti A. ogale, "A Survey of techniques for human detection from video" Department of computer science, university of Maryland, College park, MD 20742.

[8]. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 886–893

[9]. Y. Mu, S. Yan, Y. Liu, T. Huang, and B. Zhou, "Discriminative local binary patterns for human detection in personal album," in *Proc. IEEEInt Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8

[10]. Y. Mu, S. Yan, Y. Liu, T. Huang, and B. Zhou, "Discriminative local binary patterns for human detection in personal album," in *Proc. IEEEInt Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[11]. Y. Liu, S. Shan, W. Zhang, X. Chen, and W. Gao, "Granularity-tunable gradients partition (GGP) descriptors for human detection," in *Proc.IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1255– 1262.

[12]. M. Enzweiler and D. M. Gavrila, "Multilevel mixture-of-experts framework for pedestrian classification," *IEEE Trans. Image Process.*, vol. 20, no. 10, pp. 2967–2979, Oct. 2011.

[13]. T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Object recognition with cortex-like mechanisms," *IEEE Trans. Pattern Anal.Mach. Intell.*, vol. 29, no. 3, pp. 411–426, Mar. 2007

[14]. S. Maji, A. C. Berg, and J. Malik, "Classification using intersection kernel support vector machines is efficient," in *Proc. IEEE Int. Conf.Comput. Vis. Pattern Recognit.*, Jun. 2008, pp.1–8.

[15]. C. H. Lampert, "An efficient divide-and-conquer cascade for nonlinear object detection," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1022–1029.

[16]. Q. Ye, J. Jiao, and B. Zhang, "Fast pedestrian detection with multi-scale orientation features and two-stage classifiers," in *Proc. IEEE 17th Int.Conf. Image Process.*, Sep. 2010, pp. 881–884.

[17]. Shaopeng Tang, Satoshi Goto, "Histogram of templates for pedestrian detection" in ICICE Trans. Fundamentals /Commun. /Electron. /Inf./&Sysst., Vol. E85 – A//B/C/D, No.xx January 20xx.

[18]. S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, and K. R. K. Murthy, "A fast iterative nearest point algorithm for support vector machine classifier design," *IEEE Trans. Neural Netw.*, vol. 11, no. 1, pp. 124–136, Jan. 2000.

Feed-forward and Feedback Timing Recovery Algorithms for MSK

[1]Supriya V
[2]M Desanna
[3] Prof. K N Pushpalatha
[1,3] Department of Electronics and Communication, Dayanandasagar College of Engineering, Bangalore
[3] Sr. Research Staff, Central Research Laboratory, BEL, Bangalore

Abstract - This research paper addresses feed-forward and feedback timing recovery algorithms for MSK. Both the algorithms do not require previous or simultaneous acquisition of carrier phase. Timing error estimation is done over AWGN channel. Performance analysis is made based on the simulation results.

I. INTRODUCTION

MSK modulation is kind of continuous phase modulation which has attractive properties such as continuous phase at bit transitions, constant envelope. MSK also makes use of the available bandwidth efficiently which is one of the main reasons why MSK has attained considerable attention in the modern wireless communication.

Timing recovery is the first synchronization operation processed by the digital receiver and so is a vital part of any synchronous receiver [1].To demodulate the signal correctly in the receiver, knowledge of carrier phase, symbol timing, and frequency offset are required [2].One of the global synchronization approaches such as maximum-likelihood approach is not very practical due to its computational complexity. There are other data-aided algorithms which are proposed to extract the timing information in [3].The timing information in [3] is extracted by the argument difference between every symbol. This algorithm however works well at high SNR cases but degrades dramatically at low SNRs.

In this paper feed-forward and feedback methods for timing recovery has been presented. Feed-forward structure can quickly capture the information, suitable for emergency communication but feedback structure can more accurately restore the signal [4]. The feed-forward method proposed is computationally less complex and shows a better performance at lower SNRs when compared to the algorithm proposed in [3]. The feedback method presented in this paper is a simple structure which is suited for digital implementation. Both feed-forward and feedback algorithms proposed in this paper do not require previous or simultaneous acquisition of carrier phase. From the simulation results it is seen that feedback method shows better performance than feed-forward method over AWGN channel.

The paper is organized as follows. In Section II the feed-forward timing synchronization is presented. In Section III feedback timing synchronization technique is presented. In
section IV simulation results are presented .In Section V conclusions are drawn.

II. FEED-FORWARD METHOD FOR TIMING RECOVERY

The feed-forward algorithm presented in this section is a data-aided algorithm based on two statistical variables. The signal model for MSK system is shown in Fig 1.

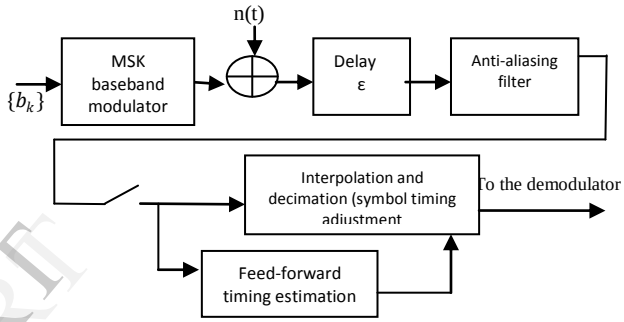


Fig 1. Signal Model for MSK System

The received signal is oversampled at the time $\left(k+\frac{i}{N}\right)T$ is written as,

$$z_{k,i} = e^{j\left(\phi(KT+\frac{iT}{N}-\varepsilon T)+2\pi f_{\Delta}\left(k+\frac{i}{N}\right)T+\psi\right)} + n_{k,i} \quad (1)$$

where ε (-0.5,0.5) is the fraction of symbol duration by which received signal is time shifted w.r.t the original signal.f_{Δ} is the frequency offset between transmitter and receiver, ψ is the initial offset and ϕ is the information bearing phase.

Timing Error Estimation

To estimate the timing error, the m-lag fourth order non-linear transformation is chosen:

$$R_m(i) = \left\{E\left(z_{k,i}\, z^*_{k-1,i}\right)\left(z_{k-m,i}\, z^*_{k-m-1,i}\right)\right\} \quad (2)$$

Due to the limited length of pilot symbols, the fourth order expectation $R_m(i)$ is obtained by averaging the samples,

$$\hat{R}_m(i) = \frac{1}{L-m-1}\sum_{k=m+2}^{L}\left(z_{k,i}\, z^*_{k-1,i}\right)\left(z_{k-m,i} z^*_{k-m-1,i}\right) \quad (3)$$

The estimated timing error is obtained by ,

$$\hat{\varepsilon} = \frac{1}{2\pi} arg\left[\sum_{index=0}^{N-1} -sgn_{index}\, FR(index)\right] \quad (4)$$

where,

$$sgn_{index} = \begin{cases} 1 & index = 1 \\ -1 & index = N-1 \\ 0 & otherwise \end{cases}$$

$FR_m(n)$ is the discrete fourier transformation of $\hat{R}_m(i)$.

III. FEEDBACK METHOD FOR TIMING RECOVERY

In this section a data-aided simple feedback structure is presented. The block diagram of the structure is shown in figure 2.

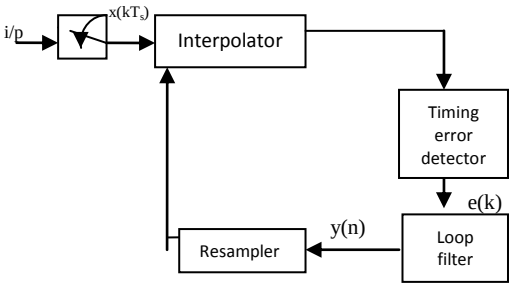


Fig (1):Feedback structure

The MSK signal model is

$$x(t) = e^{j(2\pi f_0 t+\theta)}\sqrt{\frac{2E_s}{T}}e^{j\psi(t-\tau,\alpha)} \quad (1)$$

We assume that the sampling time is $t_m = kT + nT_s + \tau_k$ where T is the symbol period, T_s is the sampling period and $T_s = \frac{T}{N}$, $k = int\left(\frac{m}{N}\right)$, $n = m\ mod\ N$ and τ_k is the timing error. We take the samples one step before and after the generic symbol interval, $kT + \tau_k$ which gives,

$$e(k) = (-1)^{D+1}\Re\{x^2(KT - T_s + \tau_{k-1})x^{*2}[(k-D)T - T_s + \tau_{k-D-1})\} - (-1)^{D+1}\Re\{x^2(KT + T_s + \tau_k)x^{*2}[(k-D)T + T_s + \tau_{k-D})\} \quad (2)$$

D is the design parameter taking integer and positive values. D=1 is good choice for MSK. The error signal from the timing error detector is sent to the loop filter. The error signal generated by the error detector is the noisy estimate of phase error. The loop filter processes e(k) in order to generate useful error by suppressing the effect of noise as much as possible. Loop filter design is done by taking the ideal filter points, the discrete time domain loop filter of the recursive equation:

$$y(n) = y(n-1) + c2*[e(n)-e(n-1)]+c1*e(n) \quad (3)$$

where $c1 = \frac{2\omega_n\xi}{K}$, $c2 = \frac{\omega_n^2}{K}$, ω_n denotes loop bandwidth, ξ denotes damping and K is the loop gain. Now we resample the resulting signal at time instants $t = kT_i$ where T_i is synchronized with the signal symbols. Now the correct set of signal samples is identified by the base point index m_k and correct set of filter samples is identified by the fractional interval μ_k[5].Here we interpolate the resulting signal from the resampler by the scheme pointed by M.Moeneclaey where the use of NCO is eliminated. Two successive interpolations are performed for time instants

$$kT_i = (m_k + \mu_k)T_s \quad (4)$$

$$(k+1)T_i = (m_{k+1} + \mu_{k+1})T_s \quad (5)$$

Subtracting these two expressions and rearranging slightly gives recursion

$$m_{k+1} = m_k + \frac{T_i}{T_s} + \mu_k - \mu_{k+1} \quad (6)$$

m_{k+1} is an integer, Then since $0 \leq \mu_{k+1} \leq 1$,

$$m_{k+1} + \mu_{k+1} = m_k + \frac{T_i}{T_s} + \mu_k < m_{k+2} \quad (7)$$

Hence the increment in the sample count from one interpolation to the next is

$$m_{k+1} - m_k = int\left[\frac{T_i}{T_s} + \mu_k\right] \quad (8)$$

To compute the fractional interval μ_k, recognize that the fractional part of the increment is zero i.e.,

$$f_p[m_{k+1} - m_k] = 0 = f_p\left[\frac{T_i}{T_s} + \mu_k - \mu_{k+1}\right] \quad (9)$$

From which we get,

$$\mu_{k+1} = \left[\mu_k + \frac{T_i}{T_s}\right] mod\ 1 \quad (10)$$

IV. Simulation Results

Feed-forward and feedback methods are simulated using Matlab. We take the oversampling factor as 8 with the length of pilot symbols L as 16.Symbol period is taken as 10^{-7}.Here the algorithm accuracy is measured in terms of variance and variance of the estimated timing error is given by,

$$var(\hat{\varepsilon}) = E\{\hat{\varepsilon}^2\}$$

$$\approx \left[\frac{1}{2\pi}\right]^2 \frac{E\{(Im\ X)^2\}}{E\{(Re\ X)^2\}}$$

where

$$X = \sum_{i=0}^{N-1} \frac{1}{L-m-1} \sum_{k=m+2}^{L} (z_{k,i}\, z^*_{k-1,i})(z_{k-m,i} z^*_{k-m-1,i}) e^{-j4\pi f_\Delta T} \left(e^{\frac{-j2\pi(N-1)i}{N}} - e^{\frac{j2\pi i}{N}}\right)$$

Figure(1) shows the timing estimation of feed-forward scheme at different SNRs.

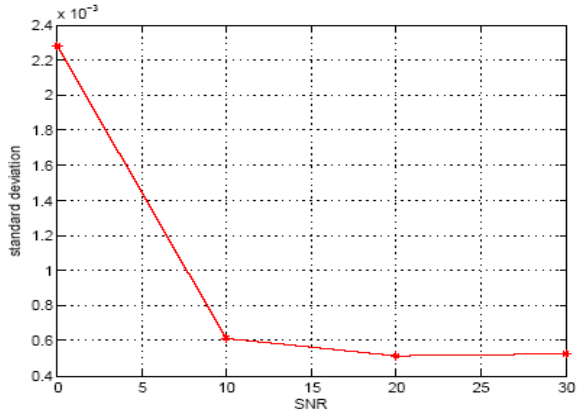


Fig (1). Standard Deviation of estimated timing error vs SNR

In the feedback method the sampling frequency f_s is taken as 64MHz with loop bandwidth is taken as 0.05. Figure(2) shows the timing error estimation of feedback loop over AWGN channel.

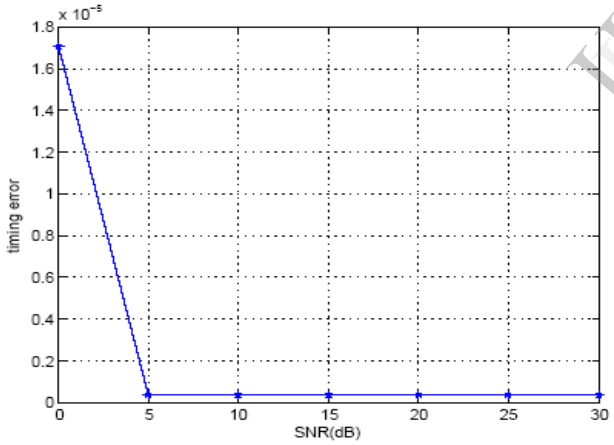


Fig (2). Estimated timing error vs SNR

From the simulation results we can infer that feedback method can more accurately restore the information in the presence of AWGN channel when compared to feed-forward method. After a certain period it is possible to make the error almost zero in feedback method so that accuracy of getting back the original signal is more when compared to feed-forward. Thus feedback method shows a better performance when compared to feed-forward scheme.

V. CONCLUSION

In this paper novel feed-forward and feedback methods for timing error estimation for MSK are proposed. Timing error for both the methods at different SNRs have been plotted and analyzed. From the simulation results it is seen that feedback method shows a better performance than feed-forward scheme and feedback technique can more accurately restore the information when compared to feed-forward scheme.

REFERENCES

[1]. Eynard, Goulven, and Christophe Laot. "Non data aided timing recovery algorithm for digital underwater communications." In *OCEANS 2007-Europe*, pp. 1-5. IEEE, 2007.

[2]. Yao, Yao, and Tung-Sang Ng. "Feed-forward carrier frequency and timing synchronization for MSK modulation." In *Proceedings of the 2003 Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing, and the Fourth Pacific Rim Conference on Multimedia*, vol. 1, pp. 549-553. 2003.

[3]. Lambrette, Uwe, and Heinrich Meyr. "Two timing recovery algorithms for MSK." In *Communications, 1994. ICC'94, SUPERCOMM/ICC'94, Conference Record,'Serving Humanity Through Communications.'IEEE International Conference on*, pp. 1155-1159. IEEE, 1994.

[4]. Mi, Zhang, Yan Xiao, and Kaiyu Qin. "MSK signal timing recovery technology research." In *Communications, Circuits and Systems (ICCCAS), 2010 International Conference on*, pp. 141-143. IEEE, 2010.

[5]. Gardner, Floyd M. "Interpolation in digital modems-Part I: Fundamentals." *IEEE Transactions on communications* 41, no. 3 (1993): 501.

Image and Video Enhancement Based On Weighting Distribution and Gamma Correction

R. Sridhar and N.P. Deepa
Department of electronics and communication, DayanandaSagar College of Engineering, Bangalore, Karnataka, India

***Abstract* -- In *this*paper, we present an efficient method to transform histogram and to improve contrast in images. Contrast enhancement is one of the most important issues in image processing, computer vision and pattern recognition. We present techniques that enhance the brightness of low contrast image via gamma correction and distribution of luminance pixels. Then, the gamma corrected frame is applied to pixel intensity transformation which can be implemented by nonlinear transfer function followed by center surround enhancement. Both pixel intensity transformation and center surround enhancement are carried out for value component and preserving spectral composition and purity of the color. On the other hand, to enhance video sequence, the proposed image enhancement uses entropy model, which calculates the temporal information with respect to differences between each frame to reduce computational complexity. Simulation results demonstrate that the proposed method produces enhanced images of comparable quality.**

Index Terms -- Contrast Enhancement, Entropy Model, Gamma Correction, Histogram Equalization, Weighting Distribution, Luminance Enhancement.

I. INTRODUCTION

Image enhancement is a process of improving visual appearance of an image to make it more acceptable to human or machine. Image enhancement is used as a pre-processing step in image/video processing application, medical image analysis [1]. Image enhancement is done by changing some feature of the image. Different techniques are available for image enhancement.

Contrast enhancement is the most popular and plays very important role in image enhancement techniques. Contrast enhancement will be used to perform adjustment on darkness and lightness of an image. It is mainly used to bring out the features hidden in an image or to increase the contrast of low contrast image.

Contrast enhancement is one of the most important issues of image processing applications such has image analysis, remote sensing digital photography, LCD display processing and scientific utilization. Contrast enhancement is applied only when the image/video is suffering from poor contrast due to lack of operator expertise, poor quality of capture device and the adverse environmental conditions at the time of acquisition. This results in underutilization of the dynamic range. As a result such images and videos do not reveal all the information in the captured scene. Contrast enhancement targets to eliminate these problems and thereby to obtain a more visually pleasing images and videos [2]. Contrast enhancement is an essential factor in case of dimmed images or dimmed videos.

Contrast enhancement is an essential factor in case of dimmed images or dimmed videos. The most commonly used techniques for contrast enhancement falls under two categories direct methods and indirect methods. Direct method defines criterion of contrast measurement and enhance the image by improving the contrast measure. Indirect methods, improve the contrast through exploiting the under-utilized regions of the dynamic range without defining a specific contrast term. Most methods in the literature fall into second group. Indirect methods can be further divided into following sub groups: i) histogram modification techniques. ii) Techniques that decompose an image into high and low frequency signals for manipulation, and iii) transformed based techniques.

Among these three subgroups histogram modifications techniques receives the most attention due to its straightforward, easy and fast implementation. Through histogram modification, the original gray level is assigned a new value. As a result. The intensity span of the pixels is expanded. Histogram modification technique onlystretches the distribution of the intensity. Many contrast enhancement techniques have been used in order to optimize the visual quality of the image for human or machine vision through grayscale or histogram modifications. All these methods try to enhance the contrast of input image. [3]

Video enhancement is one of the most important and difficult components in video research. The main aim of

the video enhancement is to improve the visual appearance of the video. Video enhancement plays an important role in analysis, recognition, surveillance, traffic, criminal justice system, detection, segmentation.

The rest of the paper is organized as follows: section II provides a brief discussion of related works. Section III presents our proposed method in detail. In section IV, the simulation results. Finally, conclusion and future work is presented in Section V.

II. PREVIOUSWORKS

Histogram is defined as the probability distribution of each gray level in a digital image. Histogram equalization is one of the well-known methods for enhancing the contrast of given images. Histogram equalization is an effective technique to transform a narrow histogram by spreading the gray level clusters in the histogram and is adaptive since it is based on the histogram of given image. However histogram equalization is rarely employed because it may significantly change the brightness of an input image and cause undesirable results. In order to solve the aforementioned problem associated with histogram equalization, many variants of histogram equalization that preserves the image brightness have been proposed. Kim proposed Brightness preserving Bi histogram equalization (BBHE) to overcome the problem. BBHE first separates the input images histogram into two sub histogram by it mean. Next it equalizes the two sub histograms separately.. Thus BBHE can preserve original brightness to certain extent [4].

Wan, Chen and Zhang Proposed Dualistic Sub Image Histogram Equalization (DSIHE), which is similar to BBHE except the threshold for histogram segmentation is not the mean of the input image, but the median of the input image brightness. It produces good image enhancement, but equalization effect is reduced. [5].

Chen and Ramli introduced Minimum Mean Brightness Error Bi Histogram Equalization (MMBEHE), which is the extension of BBHE, MMBEHE performs separation based on threshold level, which would yield minimum difference between input and output mean. [6].

Chen and Ramli also proposed Recursive Mean Separate Histogram Equalization (RMSHE), it recursively separates the input image histogram into multiple sub histograms. If the sub histogram is too large then no significant enhancement is performed [7].

Sim, Tso, Tan proposed recursive sub image histogram equalization (RSIHE), which has multiple local intensities to overcome the drawback of DSIHE. Instead of separating image once, it recursively separates into multiple sub histograms. This method has good contrast enhancement effect [8].

Kim and Chung introduced recursively separated and weighted histogram equalization (RSWHE), to enhance the image contrast as well as preserve the image brightness. RSWHE consists of three modules: histogram segmentation, histogram weighting and histogram equalization. The histogram segmentation splits the input histogram into two or more sub histograms recursively. Histogram weighting changesthe sub histograms trough weighting process based on normalized power law function. Histogram equalization equalizes the weighted sub histograms independently. [9].

Contrast enhancement can be optimized by the histogram modification framework, which incorporates penalty terms for histogram deviation as well as minimizes a cost to compute a target histogram. In order to accurately preserve brightness, the automatic weighting mean separated histogram equalization (AWMHE), method uses two modules. Automatic histogram separation to separate the input image histogram recursively based on weighting mean function. The piece wise transformation function equalizes the sub histograms in small scale details able to achieve contrast enhancement. This technique uses only one dimensional histogram, even if it might possess spikes which compress other gray levels for distribution [10]. To overcome the previous discussed problem, the two dimensional histogram is used to generate contextual and variational information (CVC) in the image while the Gaussian mixture model (GMM) can also be used to compensate for gray level distribution of the image. The contextual and variational contrast enhancement method is more effective at showing the visual quality of the image, because it directly constructs a prior probability, which further represents information of the image. However the CVC method requires a high level of computation when increasing the gray level differences between neighbouring pixels [11].

III. PROPOSED METHOD

In order to compensate for the above limitations of these methods, a technique must be developed which creates a balance between visual quality and low computational costs. In this paper, two methods are proposed to accomplish this goal. a) Gamma correction with weighting distribution technique (GCWD) for image enhancement. b) Temporal Based (TB) method for video enhancement.

Fig.1 shows the block diagram of proposed GCWD method. A low contrast image is used as an input; most of the pixels are densely distributed in the low levelregions. In histogram analysis, the input image is converted into grayscale imageand histogram equalization is performed to change the value in intensity image. The weighting distribution function is applied to slightly modify the statistical histogramand lessen the generation of adverse effect. Thus, the fluctuant phenomenon can be smoothed.Gamma correction reduces the over enhancement by preserving the brightness of the color image.

The gamma correction parameter can be formulated by basic form of power law transformation function defined as [12].

$$S = L_{max} * (L / L_{max})^{\gamma} \quad (1)$$

Where γ are positive constants
$L_{max}$is the maximum intensity of the input. The intensity L of each pixel in the input image is transformed as S after performing Eq. (1).

The exponent in the power law equation is referred to as gamma (γ). The process which is used to correct this power law response is called gamma correction. If the γ>1 then image appears as a darker image and if γ<1 leads to contrast stretching, so images appears as brighter thus having opposite effect. Thus γ>1 have exactly opposite effect as those generated with γ<1.

In the next step, gamma corrected image applied to the pixel intensity transformation using nonlinear transfer function followed by center surround enhancement [13].

The luminance enhancement through pixel intensity transformation implemented by using nonlinear transfer function which is defined as

$$V_{LE} = \frac{V^{(0.75z+0.25)} + 0.4(1-z)(1-V) + V^{(2-z)}}{2} \quad (2)$$

Wherez is the dependent parameter, defined as,

$$Z = \begin{cases} 0 \text{ for } L \geq 50 \\ \frac{L-50}{100} \text{ for } 50 < L \leq 150 \\ 1 \text{ for } > 150 \end{cases} \quad (3)$$

L is the intensity levels

The above transfer function can largely increase the luminance of those dark pixels while brighter regions have lower negative enhancement.

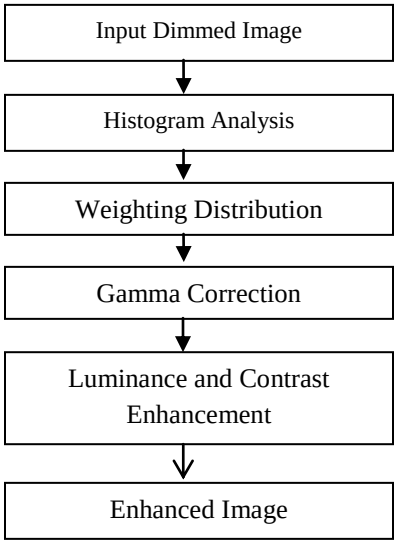


Fig.1. Block Diagram of GCWD Method.

2D discrete convolution is carried out on the original V channel image by using Gaussian function.

The following equation defines the process of contrast enhancement for V channel.

$$V_{CE}(i, j) = V_T(i, j)^{E(i, j)} \quad (4)$$

Where E (i, j) is defined by

$$E(i, j) = R(i, j)^G = [\, V_T(i, j) / V_F(i, j) \,]^G \quad (5)$$

V_{CE} (i, j) is the contrast enhanced V channel image.R (i, j) is the ratio between Gaussian filtered and original value component image. G is the image dependent parameter determined by using standard deviation of input value channel image V_F.

According to the analysis, by applying HSV color model which approximately equal to RGB model we can enhance the color image to be acceptable to human vision. In HSV color model, hue (H) is identical to dominant wavelength which represents the spectral composition of colors. Saturation (S) it is equivalent to purity which represents the purity of the color. Value (V) it is also called luminance which represents the brightness or intensity of an image. The color image can be enhanced by preserving H and S while enhancing only V.

Hence ProposedGamma Correction with Weighting Distribution (GCWD) method was applied to V component for color contrast enhancement.

The GCWD method can progressively increases the low intensity and avoid the significant decrement of the high intensity. Thus the proposed GCWD method can enhance a color image without generating artifacts or distorting the color.

In addition to the image contrast enhancement, we also propose a temporal based method (TB) technique to further reduce the computational complexity required by GCWD method to enhance video sequence.

The flow chart of temporal based method applied for video contrast enhancement is as show in Fig. 2. At the beginning of the process, the first incoming frame is directly stored in frame storage, which is used to generate mapping curve for the proposed GCWD method. For subsequent incoming frame, the entropy model can be used to measure the information content between two successive frames. The information content of each frame is approximated by the following entropy formula:

$$H = -\text{sum}(P.* \log2(P)) \quad (6)$$

Where P is the number of histogram count.

When the absolute difference between the current H and previous H exceeds threshold T_h, the frame storage can be updated by incoming frame, while the transformation curve is also modified. In this situation, T_h is empirically set to 0.05. Otherwise, the existing mapping curve is directly applied to transform each intensity level in the incoming frame.

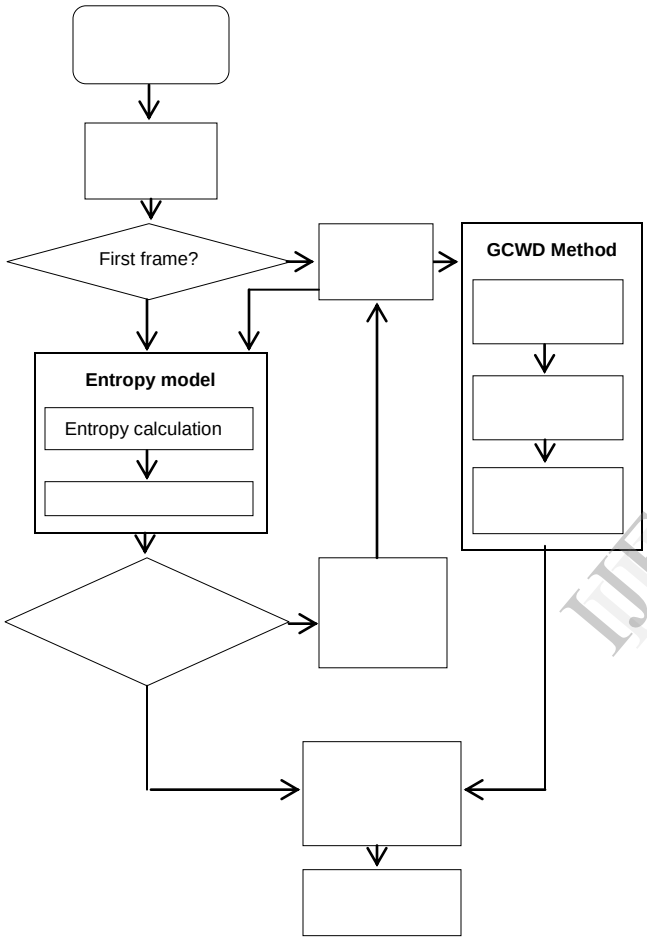


Fig. 3.Flow chart of the temporal based method

IV. Simulation Results

This section summarizes the simulated results produced by both GCWD method and TB method. For image enhancement GCWD method was applied and TB method was applied to video enhancement.

In general, change in the contrast can be caused by many factors which are common to outdoor scenes, such as intensityof the sunshine, the location of light source, cloud cover. Similarly the image of indoor scenes, the quality is often affected by interior lighting.

A. Image Enhancement.

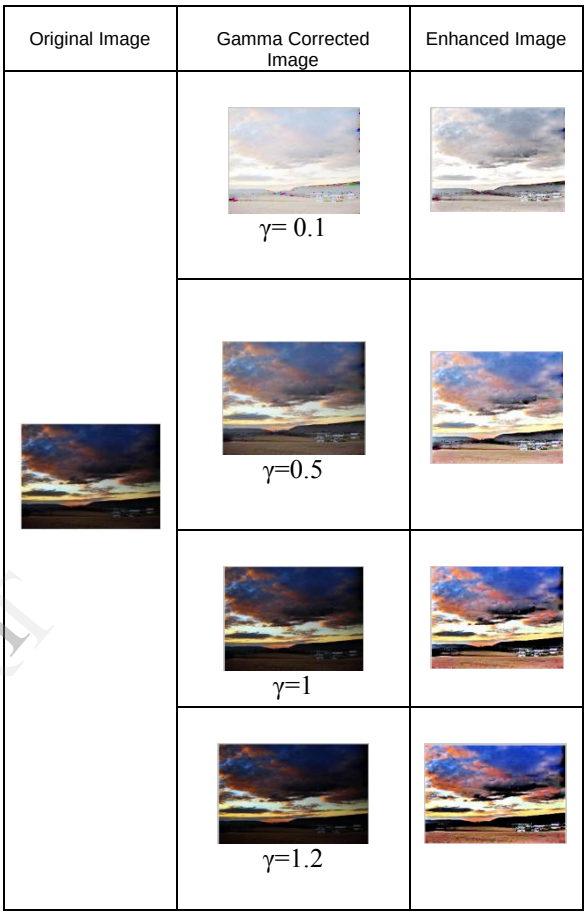


Fig. 4.Image enhancement for different values of gamma and corresponding enhanced image.

Contrast enhancement for dimmed images performed by proposed method for the above image in Fig. 4 shows original image, with its gamma corrected image and there equivalent enhanced image

B. Enhancement of Video Sequence.

In addition to the image enhancement, video contrast enhancement is also achieved. Fig.6 shows seven sampled frames of the man walking video sequence and its enhancement results generated by TB method. In the video, the man walking in dimmed area. The method increases the contrast between the man and background without distorting color or generating artifacts.

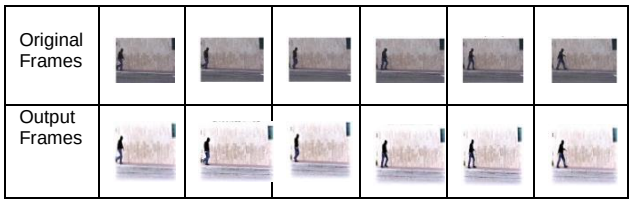


Fig.6. Seven sampled frames of the man walking and the enhancement results generated by TB method.

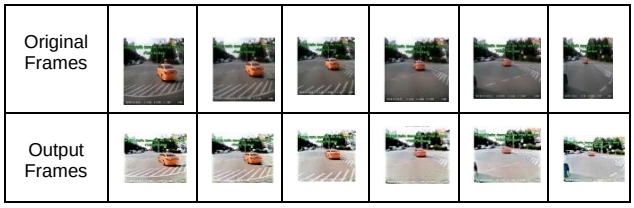


Fig. 7.Seven sampled frames of car crossing traffic signal and the corresponding enhanced frames by TB method.

Fig.7. shows the car crossing traffic signal video sequence and its enhanced results generated by TB method. This method increases contrast without distorting the color and generating artifacts.

Thus, as a result the TB method for man walking and car crossing traffic signal sequences can reduce the processing time, with dependent on temporal similarity of the sequence.

V. Conclusion And Future Work

In this paper, an enhancement method for both images and video sequences are presented. The proposed method composed of following steps. First, the histogram analysis provides spatial information of single image. In the second step, the weighting distribution is used to smooth the fluctuant phenomenon and thus to avoid generation of unfavourable artifacts. In the third step, gamma correction enhances the image contrast through smoothing curve. Finally luminance and contrast enhancement is applied for the gamma corrected frame to the value component. Furthermore, we employed temporal information to reduce the computational time for several image frames of video sequence. Based on the difference in the information content, the entropy model is used to determine whether or not frame storage should be updated. According to the simulation results indicates that, the proposed method can not only enhance image/video but also keep original image luminance without distorting the color and generating unfavourable artifacts. The proposed approach can be implemented in real time video system with limited resources.

REFERENCES

[1] Rafel C. Gonzalez, Richard E. woods, Steven L. eddins, *"Digital Image Processing using MATLAB,"* Second edition. Tata McGraw Hill Publication, 2010.

[2] T. Arici, S. Dikbas and Y. Altunbasak, " *A Histogram Modification Framework And Its Application For Image Contrast Enhancement,"IEEE trans. Image process.,vol. 18, no. 9, pp. 1921-1935, sep. 2009.*

[3] H-D Cheng and A.L Negrate, *"Contrast enhancement technique based on local detection of edges,"* compute. Vis, Graph., Image process.,Vol. 46, no. 2, pp 162-174,May 1989.

[4] Y. Kim, *" Contrast Enhancement using Brightness preserving bi histogram equalization,"IEEE Trans.* Consum Electron., vol. 43, no. 1, pp. 1-8, Feb. 1997.

[5] Y. Wan, Q. Chen, and B. Zhang, *"Image enhancement based on equal area dualistic sub image histogram equalization method,"* *IEEE Trans.* Consum. Elecro., vol.45, no. 1, pp. 68-75, Feb. 1999.

[6] S.D Chen amd A. Ramli, *"Minimum mean brightness error BI histogram equalization in contrast enhancement,"IEEE Trans.* On consumer Electronics, vol.49, no.4, pp.1310-1319, Nov.2003.

[7] S.D Chen amd A. Ramli, *"Contrast enhancement using recursive mean separated for histogram equalization for scalable brightness preservation,"IEEE Trans* on consumer Electronics, vol. 49. No .4, pp.1301-1309, 2003.

[8] K. S. Sim, C. P. Tso, and Y. Tan, *"Recursive sub image histogram equalization applied to gray scale images."*Paternrecognit. Lett. vol. 28, no.10, pp. 1209-1221, jul.2007.

[9] M. Kim and M.G. Chung, *"Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement,"IEEE Trans.* Consum.Electrin, vol.54, no.3, pp. 1389-1397, aug.2008.

[10] Fan-chieh, cheng, *"color contrast enhancement using automatic weighting mean separated histogram equalization,"* international journal of innovative computing, information and control, vol .7, no.9, 2011.

[11] T.Celik and T. Tjahjadi, *"Contextual and variational contrast enhancement,"IEEE Trans. Image process, vol. 20, no. 12, pp.3431-3441. Dec. 2011.*

[12] Shih-Chia Huang, Fan-Chieh Cheng and Yi-Sheng Chiu *"Efficient Contrast Enhancement Using Adaptive Gamma Correction with Weighting Distribution"IEEE Transaction on image processing* VOL. 22 No. 3, March 2013.

[13] Divya Devan, Sri Le Kshmi Das, Neethu S.S, Shreyas L,*"color image enhancement using non-linear transfer function and quality measurement using reduce reference metrics,"* IOSR journal of engineering., vol. 3. PP 36-43. e-ISSN: 2250-3021, p-ISSN: 2278-8719. Jan 2013.

Implementation of a Modified BB84 Algorithm for Secure Key Exchange in a Normal Network

Sandeep V,
Fouth Semester, M.Tech in Digital Communication,
Acharya Institute of Technology,
Bangalore.

Niranjan A.
Senior lecturer,
M.N.Technical institute,
Bangalore.

Abstract - The basic idea behind cryptography as a discipline was to research how valuable data & information can be protected? from unauthorized parties. Quantum cryptography is one of the recent advancements occurred within that discipline. Many research papers have been done to develop the algorithms, while others, to propose new implementations of these algorithms to tackle a specific problem. A conventional security mechanism such as Symmetric Key Encryption demands for the usage of a single key by the two parties (principals) involved in communication. Hence it is required for the parties to exchange the key in a secured manner. There are several public key algorithms such as RSA. However these algorithms are becoming untrustworthy because an attacker by using a modern technology may easily determine the contents of exchange and use this key to recover the data that is getting exchanged. Quantum cryptography involves the use of a number of Quantum Key Distribution (QKD) protocols such as BB84, B92, EPR. In this paper, an attempt is made to establish the secure communication channel by modifying the existing BB84. The proposed approach is tested on a wide range of inputs to analyse the required key size to be chosen by the user to obtain a key that is suitable for DEA, TDEA &AES.

Keywords: Symmetric Encryption, Public Key Encryption, Quantum Cryptography, QKD, BB84

1. INTRODUCTION

Cryptography generally aims at providing confidentiality, integrity and authentication services to the parties involved in communication. Symmetric Encryption and Asymmetric Encryption are two different existing encryption techniques, where Symmetric encryption uses single key by both the parties for the exchange of data while Asymmetric encryption uses two different keys called the public and private key pairs. For both techniques to succeed the key must be exchanged in a secured manner (Secure Key Distribution). The parties after agreeing upon the key use it for encryption and decryption of the data.

The QKD model provides a mechanism to exchange secret keys based on a security method that is rooted in the laws of Physics [3] [4]. Since this approach is still at its early stages, many researches and developments have been proposed and some had been implemented to improve it. Quantum cryptography uses our current knowledge of physics to develop a cryptosystem that is not able to be defeated - that is, one that is completely secure against being compromised without knowledge of the sender or the receiver of the messages. The word ***quantum*** refers to the fundamental behavior of the smallest particles of matter and energy: ***quantum theory*** explains everything that exists and nothing can be in violation of it. Quantum cryptography uses the ***Heisenberg Uncertainty Principle***, which holds that when a phenomenon is observed its characteristics are always affected by the act of observation.

1.1 Related work

Diffie and Hellman, presented a secure key agreement protocol that can be carried out over public communication channels and is still widely used. Even though the protocol seems to be quite simple, it is vulnerable to certain attacks. Diffie-Hellman key agreement protocol (DH protocol) has vulnerability is compounded by the fact that programmers often do not have a proper understanding of the security issues. Brassard and Bennett during the year 1984, presented BB84 to demonstrate an efficient procedure for safe key distribution. The main objections to QKD have been the expense.

Until now, the idea of leasing fibers from telecoms providers has put potential users off from the first hurdle. Also, BERs (bit error rates) caused by a combination of the Heisenberg Uncertainty Principle and microscopic impurities in the fiber make the system unworkable.

2. The BB84 protocol

The BB84 protocol [3] was proposed by Bennet & Brassard during 1984. This protocol makes use of four non-orthogonal polarization states (0º, 90º, 45º,135º) that are classified as

- Rectilinear states (0º, 90º) and
- Diagonal states (45º,135º).

The Rectilinear states are represented by the " +" symbol and the Diagonal states by the " X" symbol. The "+" & " X" are considered the "photon bases" and their equivalent binary values are called the "quantum bits" or "qubits" . The conventions used are listed in Table-1.

Polarization State	Photon base	Symbol	Degrees	Qubit
Rectilinear state	+	—	0º	0
		|	90º	1
Diagonal state	**X**	/	45º	0
		\	135º	1

Table 1: conventions used

The BB84 typically involves the following steps

1. a. Polarization state generation phase
   b. Selection of polarization states within the chosen polarization state (—, | , /, \)
2. Transmission of the polarization states to the Reciever.
3. Measurement of the received states with randomly chosen polarization states.
4. Valid state determination by comparing the received states with the generated states.
5. Saving the valid states and transmitting them to the sender.

Both the parties can now choose to use the qubits of the valid states as the one time session key for the secure exchange of data. Consider the following example shown in figure 1 to understand the working of BB84.

ACTIONS	Photon bases, polarization states and qubits							
Sender(S) randomly generates photon bases	+	+	X	+	X	+	X	+
S selecting the polarization states within the chosen base and transmits the bases	-	|	/	|	/	-	\	-
Reciever(R) randomly generates photon bases	X	+	+	+	X	+	X	X
R's polarization states within the chosen base	/	|	-	|	/	-	\	\
Valid data		1		1	0	0	1	
Transmitted & added to their sequences		1		1	0	0	1	

Figure 1: Working of BB84

3. PROPOSED APPROACH

Unlike the BB84 protocol which simply relies on the public channel for the exchange of polarization states, the proposed approach uses a secured channel for the exchange of this information. The sender in the proposed approach, encrypts the polarization states before transmission. The reciever decrypts the received states and follows the same procedure as that of the standard BB84 approach upto the transmission of the valid states. However instead of sending the valid states in plain text format the sender encrypts the valid states before transmitting to the Receiver. The reciever has to decrypt the incoming message and finally before sending the valid state message back to the sender it encrypts the message which is decrypted by the Sender.

The proposed approach typically involves the following steps

1. Polarization state generation phase
   a. Random generation of photon bases (+ or X)
   b. Selection of polarization states within the chosen polarization state (—, | , /, \)
2. **Encrypting the polarization states using a pre-shared secret key**
3. **Transmission of the encrypted polarization states to the Reciever.**
4. **Decrypting the polarization states using the sameshared key**
5. Measurement of the received states with randomly chosen polarization states.
6. Valid state determination by comparing the received states with the generated states.
7. Saving the valid states.
8. **Encrypt the valid states using the pre-shared secret key.**
9. **Transmit back the resultant message**
10. The Sender recovers the message using the same pre-shared secret key and stores the qubits for the future use as a secret key.

Additional steps included into the proposed modified BB84 are steps 2,3,4,8 & 9. These steps are highlighted for convienience. As the proposed modified BB84 protocol [1] is adopting encrypted communication, the fear of Beam Splitting, Intercept/Resend may be avoided[2].

4. RESULTS AND ANALYSIS

We developed a simulation software providing a thorough simulation of the working of a modified BB84 algorithm. Our software had two modules Secure QKD Sender & Secure QKD Receiver that were necessarily implemented on two heterogenous systems with different Operating systems.They were interconnected via a wireless router to examine its performance in terms of speed and security.
The software was tested on a wide range of inputs to analyse the required key size to be chosen by the user to be able to obtain a key for the use with different algorithms such as DEA, TDEA, IDEA, AES etc.,.

After executing the application several times with different key lengths results were obtained as shown in figure 2.

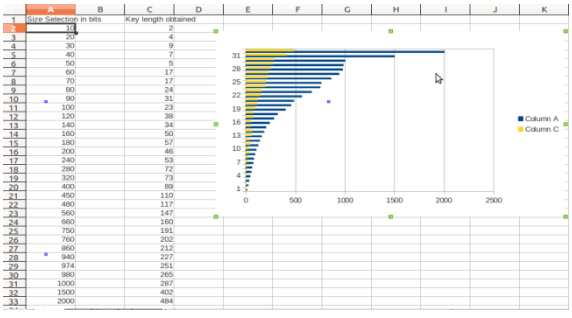


Figure 2 : A run of the application with different key sizes

Inspired by these findings we decided to show the range of key size that could be chosen to be used along with DES or TDEA or AES. These findings are shown in
figure 3.

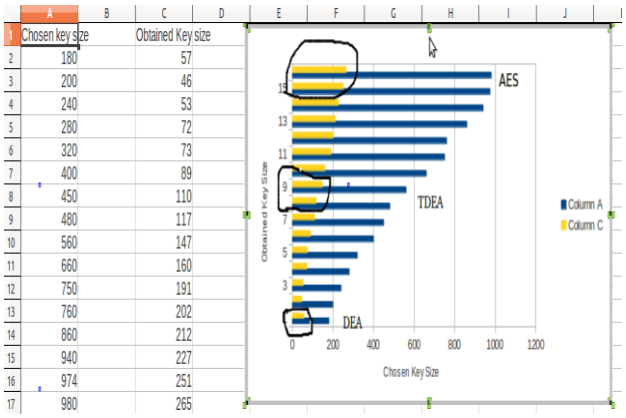


Figure 3: Key Size to be chosen to use with different Algorithms

As seen from the chart above, the user can choose a range of 180-320 to be used with DEA. A range of 400-560 for TDEAand a range of 940-980 is suitable for AES.As the key size obtained is of variable length it may be required to perform a sort of bit stuffing to be able to used for DEA/TDEA/AES.
The modified BB84 protocol combined with Symmetric encryption technology for the exchange of the qubit information provides added security.

4.1 Concerns related to Security

1. The key that is initially used for the Encryption and Decryption must be kept a secret.
2. The communicating parties can also keep the algorithm name a secret to prevent the intruder from trying to recover the key string that will be used for next encryption and decryption.

5. CONCLUSION

A simulation model and also an efficient key management scheme has been proposed for a normal network. This scheme involves a simple Key generation and Key distribution mechanisms influenced by the BB84 protocol while making it possible to be used in a non-optic normal network. It greatly reduces the communication and computation overheads of key setup involved as in Diffie-Hellman Key Distribution Algorithm and RSA. A Symmetric algorithm such as DES was used for the encryption and Decryption of the information that is exchanged between the Sender and the Receiver along with BB84 to further improve the key management scheme for generation and distribution of secret keys. The keys that are exchanged in this manner can be used along with DES, TDEA or AES based on the requirement.

6. REFERENCES

1. "A Modified QKD Approach for Secure Key Distribution Using Quantum Cryptography" Niranjan A. & C.R.Manjunath International Journal Network and Computer Engineering.ISSN 0975-6485 Volume 4, Number 1 (2013) pp. 11-15
2. Miloslav Dusek, Ondrej Haderka,Marlin Hendrych, "Generalized beam Splitting attack in Quantum Cryptography with dim coherent states" optics communication 169 (1999), 103-108.
3. C.H.Bennet & G.Brassard "Quantum Cryptograhy: Public Key Distribution & Coin Tossing"- IEEE, Bangalore, India, December 1984[pp 175-179].
4. Nicolas Gisin, Greoire Ribordy "Quantum Cryptograhy: :Riviews of Modern Physics vol-74, january 2002.
5. C.H.Bennet: "Quantum Cryptograhy using any two non orthogonal staes, physical Review letters,Vol68,pp3121-31124, May 1992.
6. Omer Abd Alkareem Jasim & Anas Ayad Abdulrazzaq "The Goals of Parity Bits in Quantum Key Distribution System"-Volume 56– No.18, October 2012
7. D. Gottesman, H. K. Lo, N. L¨utkenhaus and J . Preskill, "Security of quantum key distribution with imperfect devices", Quant. Inf. Comput. 4, pp.325–360, 2004.
8. K. J. Gordon, V. Fernandez, G. S. Buller, I. Rech, S. D. Cova and P. D. Townsend, "Quantum key distribution system clocked at 2 GHz," Opt. Express 13, pp.3015–3020, 2005.
9. William Stallings, "Network Security Essentials and Standards", Peason education,2000.
10. "A note on Quantum Cryptography", ISSN : 0975-3397 Vol. 4 No. 09 Sep 2012.

Low Complexity Modified Turbo Decoders

A.Sagar, A.Rajagopal, K.Karibasappa
Dept of ECE, Dayananda Sagar College of Engineering,
Bangalore, Karnataka-560078, India

Abstract **-- In Forward Error Correction (FEC) it is desirable to have low Bit Error Rates (BER) and low decoder complexity for reliable data transmission in Digital Communication System. FEC such as Block codes and Convolutional codes are included in standards for 2G wireless networks. Convolutional codes give greater performance compared to Block codes with increase in implementation complexity. Recently, Turbo Convolutional Codes (TCC) have been introduced in standards for 3G wireless networks. TCC are more powerful FEC codes compared to Block codes and have performance nearer to the Shannon channel capacity. TCC decoders is based on iterative decoding thus it requires more computations inorder to decode information therefore increasing the implementation complexity. These papers present a review on a new form of Low complexity decoders for modified turbo codes. Low complexity Modified turbo decoders reqiures less decoding computation compared to TCC as they use multiple concatenations of simple block codes and convolutional codes inorder to simplify decoder complexity and computation.**

Index terms **-- ILCHTC, Iterative decoding, low complexity decoding, Turbo codes, Zig-Zag codes.**

I. INTRODUCTION

In Digital communication system, Forward Error Correction plays an important role in effective usage of available bandwidth and transmission power. Shannon proved that the improvement of error-correction techniques with increasing coding gain has a limitation on the channel capacity. [1] Since then, FEC code designers are working on constructing new codes that can operate closer to the Shannon limit. However, any improvement in coding gain comes at the expense of decoder complexity and it is neccassary that practical implementation of the designed codes is compactible for existing technologies. A new class of binary parallel concatenated Recursive Systematic Convolutional (RSC) codes called turbo codes as the performance very nearer to theoretical bounds and is capable of achieving power efficiency close to the Shannon limit. Turbo codes have been adopted by the International Telecommunication Union (ITU) to effectively improve system capacity for Third-Generation (3G) wireless highspeed data services high- speed data services (CDMA2000 and W-CDMA)[1]. The error correcting capability of Turbo codes increases with increase in constraint length, however computational complexity of Turbo codes increases exponentially with the increase in constraint length of constituent convolutional codes.

A In past few years, several Techniques to achieve low complexity Turbo decoders designs have appeared in the literature and it is seen that Modified Turbo Code (MTC) provides a good tradeoff between reduced complexity and error performance. It has been shown that low complexity modified Turbo codes provide error correcting capability which is equivalent to Turbo codes. Recently, a class of modified Turbo codes termed as Low Complexity Hybrid Turbo Codes (LCHTC) has been proposed [5]. In order to improve the speed of error convergence, a new code, called as Improved Low Complexity Hybrid Codes (ILCHTC) is constructed by modifying the structure of LCHTC encoder.

Generally TCC decoding Algorithm is based on *A Posteriori Probability* (APP) which is computationally complex [1] as an alternative Maximum A Posteriori (MAP) algorithm in log domain is used then decoder complexity is about 480 Addition Equivalent Operations per Information Bit per Iteration (AEO/IB/I) and ILCHTC is multiple concatenations of simple ZigZag codes and RSC codes However, the decoding complexity of the concatenated zigzag codes is considerably lower. For instance, the MLA algorithm costs only 20 AEO/IB/Iter for rate-1/2 concatenated zigzag codes with four constituent encoders [2]. According to simulation results it is seen that ILCHTC achieve Bit Error Rate (BER) which is comparable to TCC. Simulation results show that rate-1/3 ILCHTC achieve Bit Error Rate.(BER) of 10^{-5} at Bit Energy to Noise Ratio (Eb/No) of 2.6 dB, which is 0.5 dB higher than Eb/No for TCC Moreover, ILCHTC decoder requires half the number of computations as compared to those required for TCC decoder.

The rest of the section is organized as follows, section II Describes modified turbo codes. In section III Low Complexity Modified Turbo Decoders will be described. In section IV comparsion of different decoder complexity are discussed. Matlab simulation results in section VI and conclusion in section VII

II. Description Of Modified Turbo Codes

Modified turbo codes (MTC) are a concatenation of both convolutional and block codes. ILCHTC is multiple concatenations of simple ZigZag codes and RSC codes and code rate is 1/3 as parity is generated by 2 different encoders both zigzag and Recursive Systematic Convolutional (RSC) encodes information bits.

A. Concatenated Zigzag Codes

In zigzag encoder, a sequence of N data bits is arranged in a J × K array and information bit d (j, k) is denoted as $(k + ((j - 1) \times K))^{th}$ bit. Also, d ={d(j, k)}, where, $1 \le j \le J$ and $1 \le k \le K$. Concatenated zigzag codes are parallel concatenations of M constituent zigzag codes [2]. Let zigzag parity vector of an m^{th} constituent encoder be represented by $Z^{(m)} = \{Z^{(m)}(k)\}$, where, $1 \le k \le K$ and $1 \le m \le M$. Then, zigzag parity bit is calculated progressively as follows

$$Z^{(m)}(0) = 0. \quad (1)$$

$$Z^{(m)}(k) = [\sum_{j=1}^{J} d(j,k) + Z^{(m)}(k-1)]. \quad (2)$$

Where summation symbol indicates XOR operation on Data bits. Code word, C_Z and code rate, R_Z for a concatenated zigzag code is given by

$$C_Z = \{d, z^{(1)}, z^{(1)}, \ldots, z^{(M)}\}.$$

$$R_Z = J/(J + M). \quad (3)$$

B. ILCHTC Encoder

Since RSC codes have better error-correcting capability, the Data bits are encoded by RSC code in first constituent encoder of ILCHTC [6]. To limit computational complexity only zigzag codes are used in remaining constituent encoders and ILCHTC encodes data bits array of size J × K, where N= J × K. In the first constituent encoder of ILCHTC, rate-1/2 RSC code encodes L rows of the information bit array. Let RSC parity vector for the jth row in first constituent encoder be represented by

$$r_j(1) = \{r_j(1)(k)\},\ 1 \le j \le L \text{ and } 1 \le k \le K;\ L \le J. \quad (4)$$

Then, zigzag parity bits are computed for each row of Data bit array. Fig. 1 shows General block diagram of ILCHTC encoder. Parallel concatenation of M constituent encoders forms the overall encoder. Fig. 2 shows the overall ILCHTC encoder then, transmitted code word for ILCHTC, C_{IL} is given by

$$C_{IL} = \{d, r_1^{(1)}, , r_2^{(1)}, , \ldots, r_L^{(1)}, , Z^{(1)}, Z^{(2)}, , \ , , Z^{(M)}\}. \quad (5)$$

III. Low Complexity Modified Turbo Decoders

Let the codeword C_{IL} be transmitted into channel using the bipolar format representing bit 0 as 1 and bit 1 as -1. Let $\tilde{C}_{IL}$ Denote noisy received code vector by receiver

For ILCHTC, code rate R_{IL} is given by

$$R_{IL} = J/(J + L + M) \quad (6)$$

Code rate of ILCHTC is adjusted by changing J, L and M.

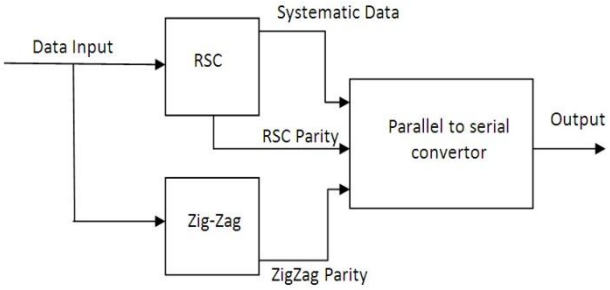


Fig 1. General block diagram of ILCHTC encoder

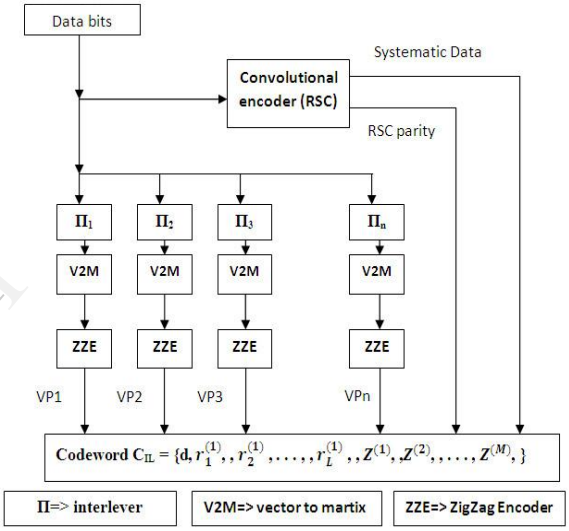


Fig 2. the overall ILCHTC encoder

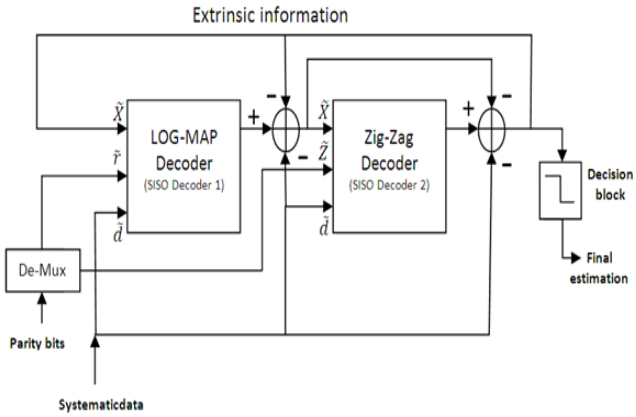


Fig 3. Low Complexity Modified Turbo Decoder

for the transmitted code word C_{IL} as $\tilde{C}_{IL} = \{\tilde{d}, \tilde{r}, \tilde{z}\}$. Here, received vectors for data bits, convolutional parity bits and zigzag parity bits are $\tilde{d}$, $\tilde{r}$, and $\tilde{z}$, respectively. Likelihood ratio (LR) of a received bit is given by

$$R(\tilde{b}\,|b) = P(\tilde{b}\,|b = +1)/(P(\tilde{b}\,|b = -1)). \quad (7)$$

Where $P(\tilde{b}\,|b = +1)$ is the probability of receiving $\tilde{b}$ if transmitted bit b is +1 [1]. Then, logarithm of likelihood ratio (LLR) of each received bit is represented by

$$b = \log(R(\tilde{b}\,|b)). \quad (8)$$

Initially, a priori value of LLR of each received bit [8] is given by

$$b = (2g/\partial^2)\,\tilde{}\,d. \quad (9)$$

For ILCHTC decoding, a priori values of LLR of received data bits $\tilde{d}$ of a codeword $\tilde{C}_{IL}$ are arranged in a J × K array.

ILCHTC decoding algorithm is as follows:

1. Decode each of L rows of the array using a priori LLRs as input to SISO convolutional decoder (Log-MAP algorithm). Output produced is a posteriori LLR of that row.

2. Taking the result of step (1) as a priori, LLRs decode each row of the array using zigzag decoder [9].

A. Max-log approximation

$$W(Z_1, Z_2, ..., Z_n) = [\prod_{j=1}^{n} \text{sign}(Z_j)] * \min 1 \le j \le n |Z_j|. \quad (10)$$

B. Forward recursion

$$F^{[q]}(p_k(i)) = \tilde{p}_k(i) + W(F^{[q]}(pk(i-1)), L_o^{[q]}(dk(i,1)), ..., L_o^{[q]}(dk(i,J))). \quad (11)$$

where i = 1, 2, . . . , I and $F^{[q]}(pk(0)) = +\infty$

C. Backward recursion:

$$B^{[q]}(p_k(i-1)) = \tilde{p}_k(i-1) + W(B^{[q]}(pk(i)), L_o^{[q]}(dk(i,1)), ..., L_o^{[q]}(dk(i,J))). \quad (12)$$

where i = I−1, ..., 2, 1 and $B^{[q]}\,\tilde{p}_k(I) = \tilde{p}(i)$

D. Extrinsic Information

$$L_e^{[q]}(d_k(i,j)) = w(F^{[q]}(pk(i-1)), L_o^{[q]}(dk(i,1)), ..., L_o^{[q]}(dk(i,J-1)) ..... , L_o^{[q]}(dk(i,J)), B^{[q]}(pk(i))) \quad (13)$$

$$L_o^{[q]}(d_k(i,j)) = \pi_k[\tilde{d}(i-1) + \sum_{k'<k}[\pi^{-1}{}_{k'}[L_e^{[q]}(d_{k'}(i,j))]] + \sum_{k'<k}[\pi^{-1}{}_{k'}[L_e^{[q-1]}(d_{k'}(i,j))]] \quad (14)$$

where Lo is initialized as an I × J matrix of zeros.

E. Final Log Likelihood Ratio Computation

$$L^{[q]}((d(i,j)) = \tilde{d}(i-1) + \sum_{k'<k}[\pi^{-1}{}_{k'}[L_e^{[q-1]}(d_{k'}(i,j))]] \quad (15)$$

For all other constituent ILCHTC decoders, only Zigzag decoder decodes the information bits. For overall ILCHTC decoders, SISO turbotype decoder structure is implemented. Iterative decoder for ILCHTC is illustrated in Fig. 3.

An iterative decoding strategy for the multidimensional concatenated ZigZag code is illustrated in Fig.4. Fig.4 (a) represents The global decoder, Fig.4 (b) consists of N decoding blocks, The block labeled by F_n consists of the interleaver π_n, the MAP or Max-Log-MAP decoder and the de-interleaver (π_n') a Its input vector of *a priori* probability ratios in the m^{th} iteration can be decomposed into two parts, denoted by $\tilde{L}_{Pn}$ and $\tilde{L}^{(m)}{}_{Dn}$ respectively. $\tilde{L}_{Pn}$ is for the parity check bits in the n^{th} dimension and it remains unchanged throughout the decoding process. $\tilde{L}^{(m)}{}_{Dn}$ is for the data bits which are updated for each iteration. The *aposteriori* ratio vector $\tilde{L}^{(m)}{}_{Dn}$ is generated by F_n, for the information bits and is delivered to the next decoding module as its input. Following the turbo decoding technique [12], the extrinsic information vector is defined by

$$W^{(m)}{}_{Dn} = L^{(m)}{}_{Dn} - \tilde{L}^{(m)}{}_{Dn} \quad (16)$$

As suggested in [l2], the extrinsic information should be prevented from circulating back to its generator. In the decoder in Fig.4(b), this is realized by subtracting the delayed values of $W^{(m)}{}_{Dn}$ front of the MAP decoding block Fn, as,

$$\tilde{L}^{(m)}{}_{Dn} = L^{(m)}{}_{Dn-1} - W^{(m-1)}{}_{Dn} \quad (17)$$

Fig.4 represents a serial updating process, i.e., the N decoding modules {Fn| n=1, 2 ... N} are activated in a serial manner. The multi-dimensional decoder in [12], whose characteristic function can be described by,

$$L^{(m)}{}_{Dn} = \tilde{L}_D + \sum_{n \ne n'} W^{(m-1)}{}_{Dn'} \quad (18)$$

Eqn. (18) can be regarded as a parallel process in which all N decoding modules can be activated concurrently during the m-th iteration, since the left hand side of (11) depends only on the results of the $(m-1)^{th}$ iteration.

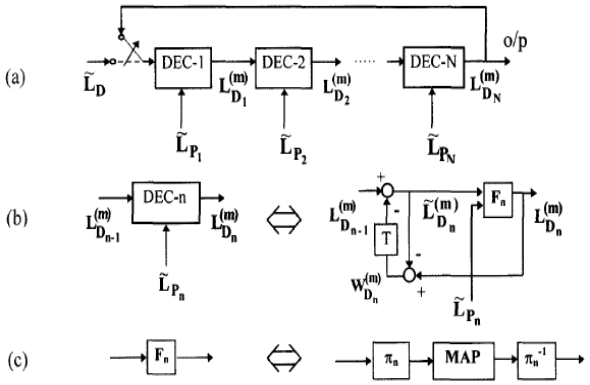


Fig 4. An N-dimensional decoder. (a) The global decoder. (b) The detailed structure of DEC-n. (c) The block labeled by Fn in (b) consists of a MAP decoder and an interleavedde-interleaver pair.

IV. Comparsion Of Different Decoder Complexity

Decoding complexity of a decoder depends upon the number of computation required to decode information bit, generally it depends on number of multiplications and additions required to decode. However Trellis length plays an important role in determine the complexity of decoders [1]. Irrespective of the code rate, the total trellis length of TCC is 2N and convolutional codes present in the ILCHTC as a Trellis length T,it is given by $T = L \times \frac{N}{J}$

Thus, trellis length increases with the increase in L.Moreover, it is maximum when L =J. ILCHTC and maximum trellis length is N. Let C_m be the number of multiplications/information bit/ iteration (M/IB/I) and C_a be the number of additions/ information bit/iteration (A/IB/I) required by a decoder. Let St be the number of states used in convolutional encoder. If Log-MAP algorithm is used to decode information bits then Cm and Ca for various codes can be found as follows:

For TCC decoder

$$C_m = 8 \times St \times M \quad (19)$$
$$C_a = (M\,(16St + 2)) - 2 \quad (20)$$

For ILCHTC decoder

$$C_m = (L(8St - 4)/J\,) \quad (21)$$
$$C_a = ((16St - 1)L/J\,) + M(5 + 4/J\,) - 1 \quad (22)$$

For zigzag decoder

$$C_m = 0 \quad (23)$$
$$C_a = M(5 + (4/J\,)) - 1 \quad (24)$$

The number of computations for each decoder is found using (19–24). Comparison of the number of computations per IB per I for each type of decoder with the number of states, St= 8 is shown in Table 1. It can be seen from Table 1 that ILCHTC and LCHTC decoders require about 50% fewer computations than that of TCC at the code rate of 1/3. Zigzag decoder requires minimum number of computations [6].

Table 1 Multiplication and Addition Complexity Of Decoders

Decoder	R	Parameters	C_m	C_a
TCC	1/2	M = 2 (punctured)	120	256
	1/3	M =2	120	256
ILCHTC	1/2	M =3, J =6, L =3	30	79
	1/3	M = 4, J =4, L =4	60	150
Zigzag	1/2	M =4, J =4	00	23
	1/3	M= 8, J =4	00	47

VI. Simulation Results

The Low Complexity Modified Turbo Decoder was simulated for a length of N=64 at Eb/no =2dB on MATLAB platform. The obtained result is shown below.

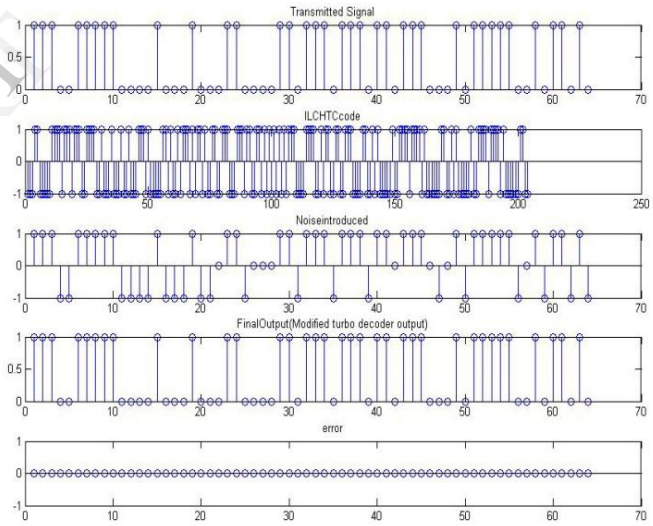


Fig 5 performance of Low Complexity Modified Turbo Decoder for N=64 at Eb/no =2dB

Fig 6 shows the MATLAB results for BER performance of TCC and ILCHTC for generator polynomial $(23, 33)_8$ with Random interleaving for length N=1024. The comparsion between the TCC and ILCHTC show that at BER of 10^{-5} is observed at 2.1dB and 2.6dB for TCC and ILCHTC respectively, where ILCHTC 0.5dB more than that of TCC.

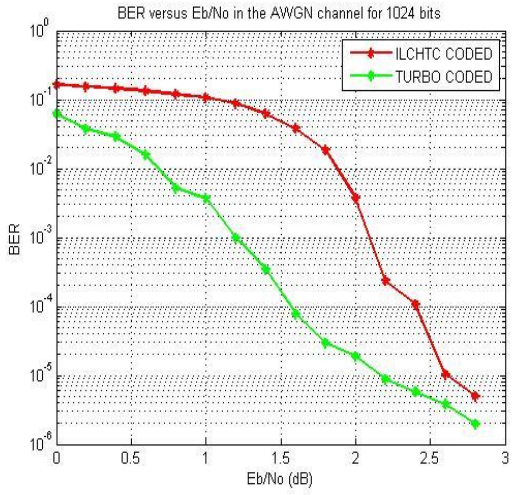


Fig 6 BER performance of Low Complexity Modified Turbo Decoder in AWGN channel

VII. CONCLUSION

In this paper a new class of modified turbo codes termed as ILCHTC and low complexity modified turbo decoders is reviewed. The ILCHTC use multiple concatenations of ZigZag codes and a convolutional code. In low complexity modified turbo decoders information bits are decoded using iterative decoding technique consisting of logmap and zigzagdecoder for better performance. The no. of computations required by ILCHTC decoder is almost 50% of the computations required by TCC decoder therefore the overall decoding complexity of low complexity modified turbo decoders is less than that of TCC.

REFERENCES

[1] M. C. Valenti and J. Sun, "The UMTS turbo code and an efficient decoder implementation suitable for software-defined radios," *International J.Wireless Inf. Netw.*, vol. 8, no. 4, pp. 203-214, Oct. 2001.

[2] Li Ping, Xiaoling Huang, and Nam Phamdo, "Zigzag Codes and Concatenated Zigzag Codes," *IEEE Trans.on Information Theory*, Vol. 47, No. 2, Feb. 2001, pp. 800- 807.

[3] Li Ping, "Turbo-SPC Codes," *IEEE Trans. On Communications*, Vol. 49, No. 5, May 2001, pp. 754-759.

[4] Al-Mohandes and M. I. Elmasry, \Design of an energy-effcient turbo decoder for 3rd generation wireless applications," in Proc. ICM 2003, Dec. 2003, Cairo, Egypt, pp. 127

[5] A. Bhise and P. D. Vyavahare, "Low complexity hybrid turbo codes," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Las Vegas, pp. 1525-1535,Mar. 2008.

[6] Archana Bhise and Prakash D. Vyavahare,"Improved Low Complexity Hybrid Turbo Codes and their Performance Analysis," *IEEE Transaction on Communications*, Vol. 58, No. 6, June, 2010, pp.1620-1622.

[7] Wade, G.: 'Coding techniques' (Palgrave Publication, India, 2004)

[8] Papaharalabos, S., Sweeney, P., Evans, B.G.: 'SISO algorithms based on Max-Log-MAP and log-MAP turbo decoding', IET Commun., 2007, 1, (1), pp. 49–54

[9] Robertson, P., Villebrun, E., Hoeher, P.: 'A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain'.Proc. IEEE ICC'95, 1995, pp. 1009–1013

[10] J.G.: 'Digital communications' (McGraw-Hill International Edition, 2001, India, 4th edn.

[11] Bhatt, Tejas; Stolpman, Victor. Structured Interleavers and Decoder Architectures for Zigzag Codes, IEEE Conference Proceedings on Signals, Systems and Computers, Oct.-Nov. 2006, pp.99-104.

[12] Ping, L., Chan, S., Yeung, K.L.: 'Iterative decoding of multidimensional

[13] Concatenated single parity check codes'. Proc. IEEE ICC,1998, pp. 131–135

Modelling Impacts of Climate Change on Road Infrastructure at Regional Scale

Pooja.V#, K C Gouda*, Ananda Kumar K R@
#M.Tech-IV Sem, Computer Engineering (Dept.of CS&E) ,S.J.B.I.T,Bangalore-560060,India
@ Department of Computer Science and Engineering, SJB Institute of Technology (VTU)
BGS Health & Education City, Kengeri, Bangalore-60, India
* CSIR Centre for Mathematical Modeling and Computer Simulation (C-MMACS)
Wind Tunnel Road, Bangalore-37, India

***Abstract* -Extreme climate events have major direct impacts on all road infrastructures. The consequences are mainly economic, but also concern safety. Climate change considerably modifies infrastructure's vulnerability to these impacts. Usually, infrastructure is designed on the basis of regulations and calculation codes which supply typical intensity values for climatic phenomena associated with a return frequency (e.g. a 10 year rainfall or a 100 year flood). While this reference event concept, based on return freq uency, has been extremely useful in the past, it is becoming dangerous lately as the underlying assumption that the climate of tomorrow will be similar to that of yesterday is no longer correct.**

According to the National Observatory of the Effects of Global Warming (ONERC), extreme meteorological phenomena will increase in number and degree in the years to come in the metro cities. In this paper we propose a prediction model to show the behaviour of road infrastructure in response to several weather related changes at a regional scale.

keywords: Weather impacts, Road infrastructure, , Spectrum sensing, wireless communication

I.INTRODUCTION

Pavement condition assessment and deterioration estimation is an integral part of all pavements and infrastructure management system. They are usually based on models which predict pavement performance based on present conditions. However, many difficulties are associated with the measurements and/or precise estimation of the inputs involved in the performance models, such as traffic flows, environmental condition etc. The uncertainty in the determination of these and other factors contribute to the difficulties encountered while developing pavement performance models. The implementing organizations have been pointing towards a need of developing an intelligent pavement performance models that can priotize pavement maintenance and rehabilitation works. Such models can forecast the pavement service life left and pavement rehabilitation needs and can help in the formulation of pavement maintenance and strengthening programs. Hence there is a need for development of performance of pavement in terms of deterioration.

The fatigue of the bituminous pavement materials under the repetitive action of traffic loading is one of the major mechanisms of structural weakening of the pavement. Since the fatigue is controlled by tensile strain which reaches maximum at the underside of the bituminous road base of a pavement under traffic loading, it is assumed that fatigue cracking would eventually initiate at the underside of the roadbase and propagates upwards till it reaches the surface. For a pavement structure containing a bituminous surfacing and a bituminous roadbase, this understanding of fatigue weakening means that the roadbase would suffer more fatigue damage than the surfacing, and the more heavily trafficked lane (the nearside lane) would suffer more cracking than the less trafficked lane (the offside lane).

We will be first analyzing the long term multi-source climate data (50-100 years) i.e temperature, rainfall, wind, humidity etc. over different geographical location (from coastal area to higher altitude area like hill station, from rural area to metro city) to study the trend pattern of the parameters. The road infrastructure data will be analyzed based on the available data (i.e soil type, road age, road type, etc.). The population data and the traffic information will be enabled to know the impact of climate on the road dynamics. All the above mentioned components will be analyzed using the different algorithms to be developed in the present work. Different data base will be developed and integrated in a cloud frame work in the high performance computing environment to analyze the big data. Finally the data mining approach, both statistical and dynamical modeling approach will be adopted for the prediction of the road behavior and the assessment of the impact of climate parameters and climate change on the transport system.

II LITERATURE SURVEY

There are different types of soil in different regions of earth. The moisture retaining property of different types of soils varies from each other. The moisture content of soil is a crucial factor which determines the pavement service life.

A. Soil type:Sandy

Sandy soil has the largest particles among the different soil types. It's dry and gritty to the touch, and because the particles have huge spaces between them, it can't hold on to water. Water drains rapidly, straight through to places where the

roots, particularly those of seedlings, cannot reach. Plants don't have a chance of using the nutrients in sandy soil more efficiently as they're swiftly carried away by the runoff. The upside to sandy soil is that it's light to work with and warms much more quickly in the spring.

B. Soil type:Silty

Silty soil has much smaller particles than sandy soil so it's smooth to the touch. When moistened, it's soapy slick. When you roll it between your fingers, dirt is left on your skin. Silty soil retains water longer, but it can't hold on to as much nutrients as you'd want it to though it's fairly fertile. Due to its moisture-retentive quality, silty soil is cold and drains poorly.

C. Soil type:Clay

Clay soil has the smallest particles among the three so it has good water storage qualities. It's sticky to the touch when wet, but smooth when dry. Due to the tiny size of its particles and its tendency to settle together, little air passes through its spaces. Because it's also slower to drain, it has a tighter hold on plant nutrients. Clay soil is thus rich in plant food for better growth. Clay soil is cold and in the spring, takes time to warm since the water within also has to warm up. The downside is that clay soil could be very heavy to work with when it gets dry. Especially during the summer months, it could turn hard and compact, making it difficult to turn. (When clay soil is worked while it's too wet though, it's prone to damage).

D. Soil type: Peaty

Peaty soil is dark brown or black in color, soft, easily compressed due to its high water content, and rich in organic matter. Peat soil started forming over 9,000 years ago, with the rapid melting of glaciers. This rapid melt drowned plants quickly and died in the process. Their decay was so slow underwater that it led to the accumulation of organic area in a concentrated spot. Although peat soil tends to be heavily saturated with water, once drained it turns into a good growing medium. In the summer though, peat could be very dry and become a fire hazard. (peat is the precursor of coal.) The most desirable quality of peat soil, however, is in its ability to hold water in during the dry months and its capacity to protect the roots from damage during very wet months. Peat contains acidic water, but growers use it to regulate soil chemistry or pH levels as well as an agent of disease control for the soil.

E. Soil type: Saline Soil

The soil in extremely dry regions is usually blackish because of its high salt content. Known as saline soil, it can cause damage to and stall plant growth, impede germination,and cause difficulties in irrigation. The salinity is due to the build up of soluble salts in the rhizosphere—high salt contents prevent water uptake by plants, leading to drought stress.

III SYSTEM MODEL

A. Existing System

Some of the earlier studies as explained in the previous section indicate there are little exploration on the combination of climate-soil-human interaction directly. In the existing system the following important points are available.

- Some factors on climate trends over regions
- Few studies on the road studies mostly on mountainous region.
- Few database but at low resolution is available

B. Existing System Disadvantages

The main disadvantages are enlisted as follows:

- The climate studies at station scale i.e over a smaller area (Taluk level) needs to be studied
- Road infrastructure and climate interaction studies are missing
- Database needs to be organized in cluster way i.e location wise, climate zone wise, rural or urban etc.
- Mathematical modeling and computer simulation approach for road infrastructure is missing.

C. Proposed System

- An integrated modeling frame work for the studies of impact of the climate change on the road at a very regional scale will be developed in the highperformance computing environment.
- Several modules of algorithms for the study and modeling aspects will be developed.
- The case studies at different locations will be implemented and the validation and the verification of the modules will be presented.
- Finally the prediction of the impact of climate change on the road will be presented using the IPCC projected climate parameters.
- A robust data base management system will be developed.
- Integration with the GIS will be incorporated.
- A GUI will be developed and implemented for the impact studies by the common users.

The required Databases are presented below:

- Climate data and climate stress factors representative of the different problems considered (e.g. road soil properties, temperature, pavement temperature, and extreme precipitations).
- Road infrastructure and the Transport information for the city and rural area(transport infrastructure, network, and transport activity).
- Physical information (e.g. sea level rise, coastal information (e.g. sea storm heights database, hydrological data, soil types.
- Engineering data and information about the underlying deterioration & damage mechanisms, maintenance practices and costs.

IV DESIGN AND METHODOLOGY OF WORK

The objective of System Design is the specification of modules and the ways these modules are to be integrated to form a complete module fulfilling its design objectives in the system design, high-level abstraction of the whole module is provided using diagrams called dataflow diagram.

A. Climate Data Collection

Several climatic parameters like temperature, Rainfall, humidity are studied. The data is collected from Indian

Meteorological Department (IMD). Several years of temperature and rainfall data are collected from IMD for the study of the changes in climate over years. Also this data is utilized to project the future climate.

B. Study of the Anthropogenic Factors

The main anthropogenic factor is the growth in population with an exponential growth in the vehicle rate. The traffic data, the rate in the increase of vehicles are studied. The major roads with heavy traffic are listed and the rates of traffic flow on these roads are noted.

C. Road Infrastructure Analysis

Several road types, their construction materials, classification based on surface typa are studied. Basically there are two types of roads which are Surfaced roads and Unsurfaced roads. In surfaced roads further classification is made based on the s urface type. It is classified into Asphalt roads, Concrete roads and Water Bound Macadem roads. Unsurfaced roads are kacha roads with mud surface. There is no surfacing done for this type of roads. Concrete roads have several advantages over Asphalt roads in terms of durability and service life. But the construction cost of concrete road is very high compared to Asphalt roads. However, it is inevitable to build concrete roads in regions with heavy rainfall to overcome the maintenance expenses.

D. Climate and Data Processing

The raw climate data of several years collected from IMD is processed using HPC systems. Firstly, the data collected is filtered and data of our study region is obtained using our algorithms. Then, the annual average and monsoon average of rainfall is calculated. Also we make an analysis of temperature and calculate the averages of annual temperature and summer temperature. Then anomaly of temperature is also plotted. The graphs for these are shown in the results and discussions chapter. Traffic data of several regions of the city is collected from various sources. The roads with high traffic are studied. Several road parameters such as longitudinal depression, pothole area, and moisture content of the soil are studied. Commercial vehicles per day, Ratio of shoulder, Rainfall, Pot hole area of the roads are studied.

E. Development of Algorithms

Several existing algorithms are analysed and their disadvantages are listed. An efficient algorithm is developed for the analysis of the multi source multi-format climate data. Also the performance parameters of the roads are calculated in the algorithm. Input parameters such as rainfall, temperature, longitudinal depression, pot hole, dry density, bearing ratio, surface cracking, and commercial vehicles per day are considered. Performance parameters such as Drainage Rating, Cracking, Roughness, Edge drop and Rut Depth are calculated.

F. System Design

The system design of the prediction model of the pavement service life is represented in fig 1. There are totally five modules in the system which takes inputs such as rainfall, temperature, traffic and road type. These modules calculate drainage rating, rut depth, edge drop, roughness and cracking which are the performance parameters of pavement service life.

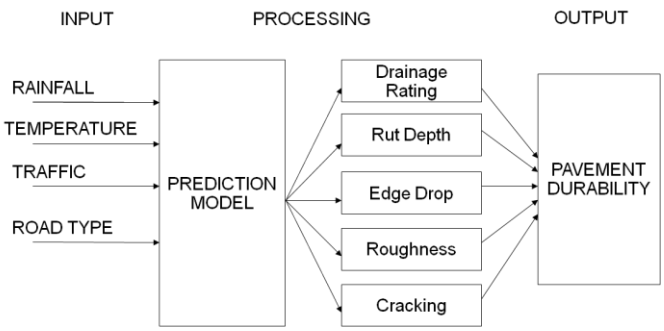


Fig.1 System Design

V RESULTS AND DISCUSSIONS

Several years of rainfall data is collected from various sources. Average of annual rainfall and monsoon rainfall is plotted. Also anomaly of rainfall is calculated and graph is plotted to determine the deviation of the rainfall from average rainfall. Temperature extremes have a lot of effects on the durability of the pavements. Temperature data is collected from various sources to analyse the annual temperature, monsoon temperature and summer temperatures. Fig 2 shows the graph plotted to show the trend of annual rainfall. Fig 3 shows the graph plotted to show the trend of annual temperature.

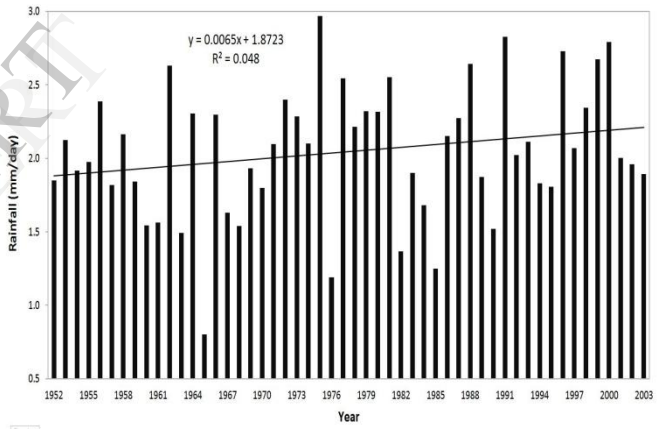


Fig. 2 Graph to show annual rainfall trend

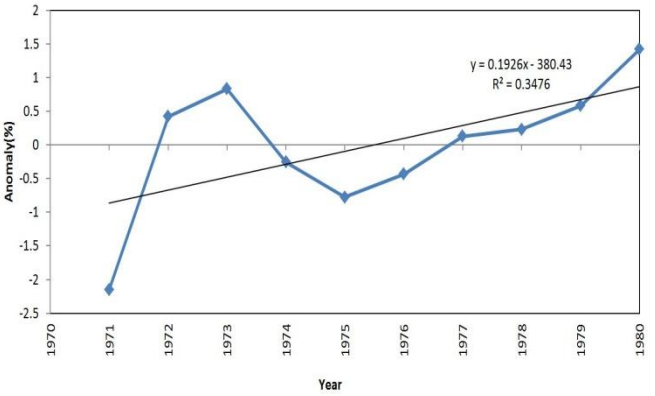


Fig. 3 Graph to show annual temperature trend

VI CONCLUSION

The main distresses identified in the rural roads are Rutting, Edge drop, Cracking and Roughness. Ravelling was found to be absent on the selected sections. A few numbers of potholes are observed on some places of the road sections due to poor drainage and construction quality. MERLIN roughness is converted into International Roughness Index in m/km. For earth, gravel, surface dressed & asphaltic concrete roads, the IRI value is minimum 2.4 to maximum 15.9. Stretches A2 & A5 is showing roughness more than higher range. Performance parameters such as drainage rating, edge drop, rut depth, roughness and cracking have been taken in the study, which mainly depends upon drainage rating, field dry density of shoulder, pothole area, California Bearing Ratio of shoulder & subgrade, edge drop, commercial vehicle per day, rainfall, surface cracking, roughness, longitudinal depression and subgrade moisture content. Validation of performance equations has also been done for collected data which gives the validity of equations. The Performance equations developed for different distress types will be used in Pavement Maintenance Management System which will be helpful in prioritization of maintenance works.

ACKNOWLEDGEMENT

The first author Pooja v acknowledges to SPARK program of CSIR C-MMACS, HOD Computer Science and Engineering & Principal, SJB Institute of Technology, Bangalore for providing the necessary infrastructures to carry out the work.

REFERENCES

1. Al-Suleiman (obedient) and Azm. S Al-Homound “A Model for Effect of Pavement Characteristics on Pavement Condition”, Journal of Indian Roads Congress, Vol. 57-1. September 1996.
2. Zi-Ping Chiang “The Study of Pavement Performance Index and Smoothness Prediction Model for Highway in Taiwan” (2000).
3. Prof. S.S. Jain, Dr. M. Parida and D. T. Thube “Optimal Strategies for Maintenance of Rural Roads in Uttaranchal” IRC International Seminar on Innovations in Construction and Maintenance of Flexible Pavements. Agra, 2-4, Sept. 2006, Pages 4-45 to 4-56.
4. R. Venkateswara Rao, C.S.R.K. Prasad “Performance Based Rural Roads Maintenance” IRC International Seminar on Innovations in Construction and Maintenance of Flexible Pavements. Agra, 2-4, sept.2006.
5. The MERLIN Low-cost Road Roughness Measuring Machine, Overseas Unit of the Transport and Road Research Laboratory, UK.

Monitoring and Propagation of Alert Information in Cloud

Unaiza Baseer
Department of Computer Science and Engineering,
Jain Global Campus, Jain University, Jakkasandra Post,
Kanakapura Taluk, Ramanagara District-562112.

Madhu B.R
Assistant Professor
Department of Computer Science and Engineering,
Jain Global Campus, Jain University, Jakkasandra Post,
Kanakapura Taluk, Ramanagara District-562112.

Abstract: **Most of the current disaster recovery techniques take time to find the failure of the server and assign an alternative or have to wait for that server to recover. In order to reduce the service time and allow uninterrupted services, in this paper the timer is used to check the status of the main server and WCF service is used to propagate alert to the virtual servers where the physical server data will be replicated. This allows virtual server to take the role of physical server and provide services to the users.**

Keywords: WCF service, Replication

1. INTRODUCTION

Cloud computing is the latest name for the emerging technology around since the mid-90s as "on-demand infrastructure". In 2000's, the automated computing became closer to what is called as "Cloud" today. From 2012 through today, for cloud computing to "cross the chasm" and attract a wider audience beyond develops and start-ups it needs to be easier, more saleable and more flexible – while being billed in a true utility model. Cloud computing 2.0 providers are offering services that are truer to the definition of cloud than ever before.

Each year, a number of natural disasters strike across the globe, killing hundreds and causing billions of dollars in property and infrastructure damage. Extreme weather events have been predicted by climate scientists and have been attributed to global warming. As number of such events increases, minimizing the impact of disasters becomes imperative in today's society. In order to overcome data loss due to disasters where in the servers get corrupted and in turn delay in giving services to the users, data backup has to be done also uninterrupted services have to be provided to the users without delay.

In the proposed approach, data is replicated in virtual servers located in different locations which act as the main server on failure of main server. And for the virtual servers to get the failure information ping command is sent to the main server for every few seconds and if the WCF service does not receive response to that command then the alert is propagated by WCF service.

2. RELATED WORK

Yoshihiro Nakajima et al. [1] designed and implemented a virtualized ICT resource management system called Management Engine. This ME with a virtualized ICT information model expresses the relationship and mapping between physical and virtual resources for carrier network services.

Manish Pokharel et al. [2] proposed cloud-based disaster recovery plan and achieved high availability, high survivability and low unavailability and low downtime with very less cost.

Katarina Grolinger et al. [3] proposed Disaster-CDM, Knowledge as a Service framework for disaster data management which stores huge amount of data while maintaining high availability using NoSQL database and various cloud solutions. Search of disaster data, interoperability, and integration were facilitated through knowledge acquisition and knowledge delivery which applies language processing, information extraction, and retrieval techniques that adds structure and metadata to largely unstructured disaster data. Knowledge delivery services integrate information from different databases and deliver knowledge to consumers. Disaster-CDM is still at the design stage, and only a part of the simulation model data acquisition process is included in the work.

Vijaykumar Javaraiah proposed a backup of data at consumer premises. With minimal cost over and above the cloud services, consumers can have peace of mind. Any solution without backup is not complete. Business continuity and disaster recovery is very much essential for any business. The lack of backup in cloud computing solutions must be plugged and with this solution, the negative impact on business can be avoided. This simple solution should address online backup, disaster recovery and also eliminate the dependency on cloud service providers.

Zhang Jian-hua and Zhang Nan [5] proposed the transformation of domestic ISP new requires including content integration, cross boundary storage, magnanimity, and centralized storage. Business diversification focused on the needs of storage shared, and especially several terminal expansions were dependent on storage. For businesses with limited resources, cloud storage appears to be a good solution.

Jianxin Li et al. [6] iROW is designed to solve some performance penalties on snapshot key operations and I/O operations. iROW uses bitmap to replace the high-cost multi-level index tree structure, which is commonly used in existing solutions; iROW combines redirect-on-write with copy-on-demand; iROWgives the disk space allocation

function back to the host machine's file system. These measures have enhanced both snapshot key operations performance and I/O performance of iROW. In addition, because of the host machine's file system supports sparse file, iROW also achieve that the VM disk image gradually increases with the actual disk usage.

Manish Pokharelet, et al [7] proposed the reliable approach in recovering from disaster where the cloud computing is used as a tool in managing the disaster in the system of organization and is analyzed using Markov model.

Timothy Wood, et al., [8] DR operational assumptions and system model are considered to recover an application to the point of crash.

Kashif Munir, et al., [9] Proposed a cloud security model with security framework which identifies the security challenges in cloud computing.

Shaftab Ahmed., et al [10] Data archiving and storage model are used which focuses on the information security issues while migrating to a cloud environment through which the confidence of end user can be gained partially with the use of cryptographic techniques.

3. PROPOSED SYSTEM

The proposed system aims at securing the data from data loss and providing uninterrupted services to the users in case of disaster.

In the proposed scheme, the user registers with the cloud service provider and access main server for storing and retrieving of the data from the cloud server. As shown in Figure 1, when the user stores the data, that data is replicated to the virtual servers located in different locations. To check the status of the main server, WCF service sends ping command to the main server every second and the main server responds to the ping command. When the main server doesn't respond to the ping command, the WCF service will send an alert to the virtual server.

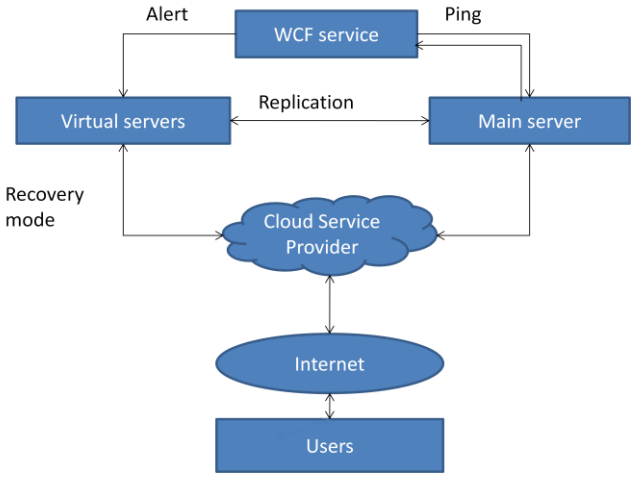
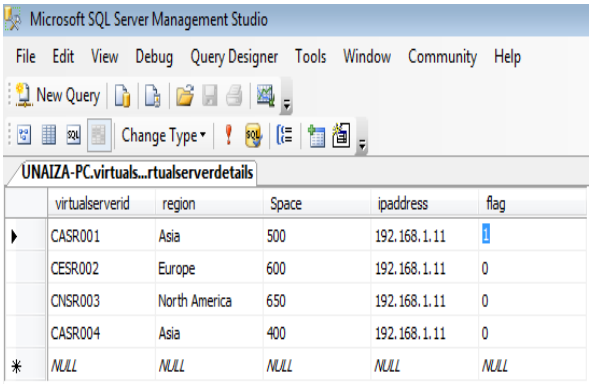


Figure 1: Monitoring and Alert propagation during disaster in Cloud Computing

On receiving an alert from the WCF, the virtual servers communicate with each other and the virtual server with the highest available space acts as the main server and provides services to the users.

To reduce the delay in giving service to the users and prevent data loss in case of a disaster, replication of the data, user details, login details are replicated to the virtual servers for service in case of the failure of main server.

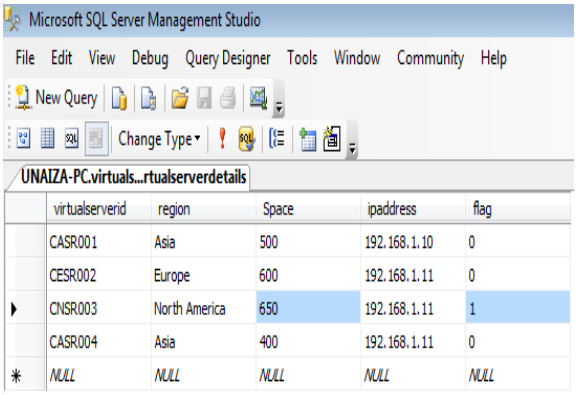
In figure 2, main server i.e. Asia as shown above is up and providing service to the users in normal conditions.



virtualserverid	region	Space	ipaddress	flag
CASR001	Asia	500	192.168.1.11	1
CESR002	Europe	600	192.168.1.11	0
CNSR003	North America	650	192.168.1.11	0
CASR004	Asia	400	192.168.1.11	0
NULL	NULL	NULL	NULL	NULL

Figure 2: Main server (Asia) is up and running (i.e., flag=1)

In figure 3, the main server is down and the virtual server located in North America is up and acting the role of main



virtualserverid	region	Space	ipaddress	flag
CASR001	Asia	500	192.168.1.10	0
CESR002	Europe	600	192.168.1.11	0
CNSR003	North America	650	192.168.1.11	1
CASR004	Asia	400	192.168.1.11	0
NULL	NULL	NULL	NULL	NULL

Figure 3: Main server is down and the virtual server located in North America is up and running which is selected based on the maximum available space (i.e., flag=1)

Server and providing service to the users which is selected based on the large available space which is as shown above.

4. CONCLUSION AND FUTURE WORK

This paper has proposed the concept of monitoring and propagation of alert during cloud disaster by which uninterrupted services to the users can be provided even in the case of a disaster. Ping command is used to check the status of the main server and WCF service is used to propagate the alert to the virtual server in case of failure. The virtual server with more available space will play the role of main server which is decided by the communication

between the virtual servers. Future work will include the replication of data to the main server and to the other virtual servers once the main server is up and ready to provide services to the requests.

REFERENCES

[1] Yoshihiro Nakajima, Hitoshi Masutani, Wenyu Shen, Hiroyuki Tanaka, Osamu Kamatani, Katsuhiro Shimano, Masaki Fukui, and Ryutaro Kawamura, "Design and Implementation Of Virtualized ICT Resource Management System for Carrier Network Services Toward Cloud Computing Era" in ITU Kaleidoscope Academic Conference, 2013.

[2] Manish Pokharel, Seulki Lee and Jong Sou Park, "Disaster Recovery for System Architecture using Cloud Computing", 10th Annual International Symposium on Applications and the Internet, 2010.

[3] Katarina Grolinger, Miriam A.M. Capretz, Emna Mezghani, Ernesto Exposito, "Knowledge as a Service Framework for Disaster Data Management", in Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, 2013.

[4] Vijaykumar Javaraiah, "Backup for Cloud and Disaster Recovery for Consumers and SMBs", Brocade Communications, 2011

[5] Zhang Jian-hua and Zhang Nan, "Cloud Computing-based Data Storage and Disaster Recovery", International Conference on Future Computer Science and Education, 2011.

[6] Jianxin L, Hanqing Liu, Lei Cui, Bo Li, Tianyu Wo, "iROW: An Efficient Live Snapshot System for Virtual Machine Disk", in IEEE 18th International Conference on Parallel and Distributed Systems, 2012.

[7] Manish Pokharel, Seulki Lee, Jong Sou Park, "Disaster Recovery for System Architecture using Cloud Computing", in 10th Annual International Symposium on Applications and the Internet, 2010.

[8] Timothy Wood, H. Andrés Lagar-Cavilla, K.K. Ramakrishnan, "PipeCloud: Using Causality to Overcome Speed-of-Light Delays in Cloud-Based Disaster Recovery", University of Massachusetts, 2013

[9] Kashif Munir, Dr. Sellapan Palaniappan, "Framework for secure cloud computing", in International Journal on Cloud Computing: Services and Architecture (IJCCSA),Vol.3, No.2, April 2013.

[10] Shaftab Ahmed, M. Yasin Akhtar Raja, "Tackling Cloud Security Issues and Forensics Model", in IEEE, 2010.

Rule Based Energy Consumption in Cluster-Based Wireless Sensor Networks

Yogavathi G
Department of Computer Science and Engineering,
Jain Global Campus, Jain University,
Jakkasandra Post, Kanakpura Taluk,
Ramanagar District-562112

Manjunath C R
Assistant Professor
Department of Computer Science and Engineering,
Jain Global Campus, Jain University,
Jakkasandra Post, Kanakapura Taluk,
Ramanagara District-562112

Abstract: **A WSN consists of interconnected spatially distributed sensor nodes without the use of wires. Sensor nodes use their communication devices in order to transmit the data being sensed to Base Station over wireless channels to other nodes in network. There is no monitoring of data entering into network during the data transmission, there are chances of over loading the network and due to which energy consumption will be increased. A Rule- based WSN system is proposed. Firstly elect Controller Node (CNode) in order to monitor the packets entering the network. Next application of rules- Interval and Jamming rule. The proposed approach helps in reducing energy consumption and enhancing network lifetime.**

Keywords: WSN, Controller Node, Rules, Interval and Jamming rule.

1. INTRODUCTION

Wireless Sensor Network (WSN) is the group of sensor nodes [1] that are distributed in a wide environment [2] to perform a specific function [1]. Sensor nodes are small devices with sensing capacity [2].An architecture of WSN includes Base Station (BS)/sink and sensor nodes. A typical sensor node consists of memory, power supply, processor, sensing, and transceiver. BS is placed close by to randomly deployed sensor nodes. Wired or wireless network can be used to connect BS to internet. BS gathers sensed data from sensor nodes and also gives instructions to them. Through internet the sensed data is easily accessible after it reaches BS. [1]

Inherent limited energy resource is one of the limitations of wireless sensor nodes [3]. Power consumption can be because of three functional domains: sensing, communication, and data processing. Energy wastage in a sensor node could be due to

- Useful or
- Waste sources

The reasons for energy waste can be either of these: firstly, idle listening; secondly collision – At the same interval if sensor node receives more than one packet that are identified to be collided. Collided packets are discarded and retransmitted which will increase the energy wastage. Third reason is overhearing – a node receives packet which was not for it. The fourth one occurs as a result of control-packet overhead- For data transmission minimum number of control packets to be used. Finally, for energy waste is over-emitting [5].

2. EXISTING SYSTEM

To deal with energy wastage there were many approaches put forth. Forming clusters were the most energy efficient approach followed to date. Grouping of associated objects into one cluster and having a cluster head to aggregate the data and forwarding of data. The cluster forwards the data without having knowledge of data which may lead to attacks and unwanted data transmissions. Though there are many approaches to overcome attacks, the methods concentrate more on attack detection but not on energy consumption and energy going waste. In [6] the author proposes an idea to monitor the data being transmitted but the approach follows data retransmission which again will consume energy.

3. PROPOSED SYSTEM

Sensing, wireless communications and computations are capabilities of wireless sensors that are deployed in an unattended environment. But the limitations of the sensor nodes such as battery, low memory and processing speed play a major role in communication. These limitations are faced because in real time approach sensor nodes are deployed randomly in hostile or impractical environment, and it becomes impossible to recharge the battery. Battery consumption can be encountered due to either normal communication or overloading of data for communication. Due to overloading or flooding of data the energy is wasted. Flooding occurs

when a node sends a packet that it received, to all its neighbors other than the neighbor which sent the packet to it. This leads to unwanted data transmission to the node which was never interested. This has many disadvantages; the most important is it is responsible for large bandwidth consumption and it wastes valuable energy.

Proposed work aims providing a ruled based energy consumption approach in the network. The approach involves two steps: Firstly, Election of Controller node after cluster formation - Monitoring node that monitors for flooding of packets. Secondly, Rule application phase - Rule are incorporated into the controller node to drop data packets which may create flood and lead to energy wastage.

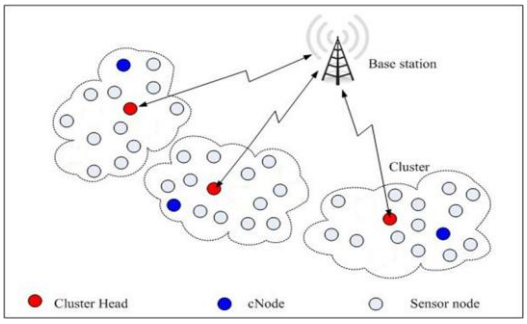


Figure 1: System architecture

Assumptions:

1) All nodes in the network are alive and posses minimum requirement of energy.
2) There is only one Base station.
3) Packets are routed via Controller Node only.

Step 1: Cluster formation

Clustering is the approach of identifying relationship among objects or grouping of similar objects. In each cluster, based on residual energy node is elected as the Controller node (CNode) and other as cluster-head (CH), while the remaining nodes are termed to be member nodes. Member nodes in their respective cluster do sense the data and forwards the same to their corresponding cluster-head or receive data from Cluster head. Cluster-heads handles the responsibilities to collect data/ transmit the data from their member nodes/BS, to aggregate them, and finally to forward the aggregated data either to neighboring cluster or to member nodes or to sink/base station. [9]

Algorithm for electing CNode:

Let RE_i be the measured residual energy of node i at time t, Let max RE be the maximum residual energy. 'M_i' be the monitor node.

1. Nodes calculate RE_i
2. If (RE_i< max RE) then
3. Node 'i' = member node
4. Else if (RE_i> max RE) then
5. Node 'i' is elected as M_i

Step 2: Rule application

There are three main phases involved: Data acquisition phase- Survey the type of data transmission and deciding on type of rules; Rule application phase- in which the pre-defined rules are applied to the data; Flooding detection phase- after rule application[10]

The rules considered for the proposed work are:

Jamming Rule: Jamming can be measured in terms of collision. If collision rate is exceeded then it is said to be creating a jam in network or collision.

Interval Rule: Interval between two consecutive messages should be within the limit. If either low or high packet interval rate then error is raised. Two attacks that can be detected by this rule are the negligence attack and the exhaustion attack. In the negligence attack, malicious node messages are not considered by the intruder. While in the exhaustion attack, the data interval rate is tampered and increments messages sending rate in order to increase the energy consumption.[7]

Algorithm for rules application:

For each data packet apply rule,

```
if (Message. Interval rate == Threshold. Interval
rate || Message. Collision rate ==
        Threshold. Collision rate)
  then "discard message"
 else
  Forward the packet to CH
end if
```

4. ANALYSIS

Proposed work aims to reduce energy consumption compared to existing techniques. It follows rule based approach in order to avoid energy wastage. Election of controller node is done to monitor the data packets. Controller node and Cluster head selection is based on residual energy calculation. The monitoring node is present in proposed work to avoid data flooding. The Controller node will be equipped with the interval rate and collision rate, which are set as thresholds. When threshold is reached, packets are dropped. This is to avoid packets that will create flooding in the network. Flooding or unwanted activity in the network leads to energy wastage. The proposed work is an attempt to avoid unwanted wastage of energy.

5. CONCLUSION

The proposed work main objective is to reduce energy consumption by clustering and rule application. Energy drop may occur due to flooding (unwanted data packet forwarding). The proposed work has two approaches built into one to avoid energy wastage. The clustering is itself followed to save energy. Rule application will avoid unwanted wastage of energy as packets creating energy drop are not entered into the network itself. Since proposed work incorporates these approaches to reduce energy consumption, it must enrich network lifetime.

6. REFERENCES

[1] Babar Nazir, Halabi Hasbullah "Energy Balanced Clustering in Wireless Sensor Network" IEEE 2012.

[2] Advanced Network Technologies Virtual Labs - IIT Kharagpur "Simulating a Wireless Sensor Network" http://virtual-labs.ac.in/cse28/ant/ant/8/theory.

[3] B. P. S. Sahoo, Satyajit Rath and Deepak Puthal "Energy Efficient Protocols for Wireless Sensor Networks: A Survey and Approach" International Journal of Computer Applications (0975 – 8887) Volume- APRIL-2012.

[4] Stefanos A. Nikolidakis , Dionisis Kandris , Dimitrios D. Vergados and Christos Douligeris : "Energy Efficient Routing in Wireless Sensor Networks through Balanced-Clustering" Algorithms 2013, 6, 29-42; doi:10.3390/a6010029.

[5] Zahra Rezaei, Shima Mobininejad "Energy Saving in Wireless Sensor Networks" International Journal of Computer Science & Engineering Survey (IJCSES) Vol.3, No.1, February 2012 DOI : 10.5121/ijcses.2012.3103 23

[6] Malek Guechari, Lynda Mokdad, Sovanna Tan "Dynamic Solution for Detecting Denial of Service Attacks in Wireless Sensor Networks" IEEE ICC 2012 - Ad-hoc and Sensor Networking Symposium

[7] Ms. Rachana Deshmukh, Ms. Rashmi Deshmukh, Prof. Manoj Sharma "Rule-Based and Cluster-Based Intrusion Detection Technique for Wireless Sensor Network", IJCSMC, Vol. 2, Issue 6, June 2013, pg.200 – 208

[8] Mamatha G, B G Premasudha "Deployment Techniques of Nodes in WSN for Multi-Domain Applications A Survey on their performance analysis" International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-3 Number-4 Issue-13 December-2013 280

[9] Zhang/RFID and Sensor Networks "Clustering in Wireless Sensor Networks" AU7777_C012 Page Proof Page 323 2009-6-24.

[10] Hosamsoleman, Ali Payandeh, Nasser Mozayyani, "Detection Collision Attacks In Wireless Sensor Network Using rule-Based Packet Flow Rate" SaeedSedighianKashi ISSN: 2248-9622 www.ijera.com Vol. 3, Issue 4, Jul-Aug 2013, pp.261-268.

Secure Multicast Scheme in Cluster Based Wireless Sensor Network

Manjunath CR[1] Maria Theresa Hoover[2], Sushma J[3], Sindhu Anand[4]
[1,2,3,4]Department of Computer Science and Engineering
School of Engineering and Technology
Jain University

***Abstract* – Wireless Sensor Networks (WSNs) are employed in numerous applications in different areas including military, ecology, and health; for example, to control of important information like the personnel position in a building, as a result, WSNs need security. Using multicast technology can significantly reduce energy consumption and prolong the life time of nodes in query-based wireless sensor networks (WSNs). Existing multicast protocols for WSNs mainly focus on covering multicast scope zone effectively, and assume secure communication between all nodes. A new efficient cluster based multicast tree (CBMT) algorithm for secure multicast Communication, in which source node uses Multicast version of Destination Sequenced Distance Vector(MDSDV) routing protocol to collects its 1 hop neighbours to form cluster and each node which have child node is elected as the Local controllers of the created clusters. It also tolerates the faults that causes due to failure of nodes.**

Index terms – CBMT(cluster based Multicast tree) cluster techniques, multicasting,security, WSN.

I.INTRODUCTION

wireless sensor network (WSN) consists of large number of sensor nodes. Each sensor node capable of sensing , data processing, computing, wireless communicating and monitoring object of interest or environmental conditions such as temperature ,sound, embedded processing, humidity, pressure, light intensity. As the technology of wireless sensor networks matures are used in numerous applications and emerging as an area of active research. Since Sensor node can be deployed in environmental monitoring, medical care, and home appliance management. It can be attacked during data transmission, It is important to provide secure communications between sensor nodes and base stations or Vice versa. Security should be considered because most of sensor networks possess various mission-critical tasks and hence WSN need security. [1][3][5][6][9][10][19]

II.SECURITY GOALS, SCHEME AND STRATEGY:

A. Security goals:

The security goals for the sensor nodes, as following:
Confidentiality: Data must be protected from being captured by any data adversaries.
Authentication: Need to know if the messages are from the node it claims to be from, determining the reliability of message's origin
Integrity: It refers to the ability to confirm the data must not been altered or changed between transmission due to environment
Availability: The network should not fail frequently

B. Security Scheme:

The security requirements of a wireless sensor network can be classified as follows[6-9,19]:
Data Authentication: Data authentication is fundamental for various applications in sensor networks and make sure that the data used in decision-making originate from the correct source is initiated from the exact source.
Data Confidentiality: Confidentiality means keeping information hidden from unauthorized party.i.eNodes should not reveal any data to unintended recipients.
Data Integrity: Data should not be changed between integrity transmissions due to the environment and make sure that any received message has not been modified sent by unauthorized parties.
Availability: Determine if a node has the ability to use the resources and the network is available for the messages making sure that the network should not fail frequently
Data Freshness: Data freshness implies that the data is recent, and it ensures that an adversary has not replayed old messages.

C. Security Strategy:

The strategy plays major role in security[8], strategy can be divided into the following several parts: node and components, connections, transmission.
Node and components: Node has two attributes: own identity and groups' identity. Own identity value can determine only one node in key establishment process, while group identity value can determine only a group.
Connection: A pair of node has multiple connection path, Direction that describes the connection path can be symmetrical or non-symmetric connection, symmetrical connection nodes can be said to be bi-direction transmission, which can make the sender and also can be the receiver; Asymmetric connection said that the direction of transmission is unique; a node only can be as the sender or recipient. Trust represents the credibility when a node completes the transmission not be captured.

Transmission: Transmission is described by brisk index which is a number between 0 and 1,the higher index the more important the information.

There are two types of techniques used to transfer the data from source to destination: Unicast routing Multicast routing

Unicast is the transferring of data from 1 node to other that is from single source to single destination and Multicast routing, which refers to the transmission of the same data to several destinations. Research shows that multicast in wireless sensor networks has more importance. It reduces the number of packet transmission, optimizes the bandwidth consumption and save the node energy. Moreover, multicast can be useful for the next generation Internet, which will integrate WSNs [20]

To prolong the life time of WSN net with limited energy resources, Multicast can better meet the requirements of network resources having the high bandwidth utilization and effective mechanisms to save energy, Multicast packets can be transmitted efficiently to reduce energy consumption effectively. However, the widely used of multicast technology in traditional wireless networks have too much difference with Wireless sensor network, such as MAODV(multicast adhoc on-demand distance vector(MAODV)),ODMRP(On-Demand Multicast Routing Protocol) and AM (adhoc multicasting) Route multicast routing protocols which have been proposed by the mobile Ad hoc networks[21]. Ad hoc network [22]is a group of computers that based on no infrastructure tries to communicate with each other. Ad -hoc networks based on their special case of usages like in battle field or relief and rescue projects; and constrains or problems with them, like battery power constrain or moving . Problem have been turned to a complex and important field.

III. MULTICASTING IN WSN

There [1][15] is a need for different methods and techniques for secure path formation. For a secure transmission, broadcasting uses the leaf nodes which are assigned keys based on all forward nodes above them. Secure multicasting scheme considers the benefits of key management techniques; the root to key management is the key distribution centre which uses a logical key. The Multicasting system provides a secure communication mechanism to ensure the data security, integrity and verifiability. Moreover, it can be justified against security attacks and known routing attacks. There are various schemes that can be incorporated to form a secure transmission path are through key management techniques or providing security to the layers or to the data that has to be transmitted. Multicast is the communication paradigm of one-to many or many-to-many, based on defined groups and constituted by members, whose interest is to receive/share the same information for a specific application. A multicast group can also have one or more senders. Multicasting in WSNs evaluates its real impact and comparing it with the conventional unicast solutions. The multicast requirement over WSNs is based on the application nature.

A. Secure Multicasting

In [12] contrast to the traditional point-to-point communication on the internet network, the major communication pattern of WSNs is multicasting. Secure multicasting pattern: while considering the benefits of a logical key hierarchy, a directed diffusion based multicast technique for WSNs. The root of key management is the key distribution centre, while the individual sensor nodes make up the leaves. Mechanisms for sensor nodes are provided by this technique by joining and leaving groups where the key hierarchy is used to effectively re-key all the nodes within the leaving nodes.

Multicast path is based on effective tree construction and hierarchical network topology in a single framework. Such integration allows the system to be optimized in terms of energy efficiency, reducing the overhead. Secure Multicast system provides a secure communication mechanism to ensure the data security, integrity and verifiability. Moreover, it can be justified against security attacks and known routing attacks.

IV. KEY MANAGEMENT AND KEY DISTRIBUTION

A. Key Management:

Key Management is the most important issue[10] in the security of Wireless Sensor Networks. It helps in maintaining the confidentiality of secret information from unauthorized users. Sometimes, it is also useful for verifying the integrity of exchanged messages and the authenticity of the sender .Since most of the public key cryptographic mechanisms are computationally intensive, most of the research studies for WSNs focus on use of symmetric key cryptographic techniques.

Key [7] [19] management is one of the most important issues of any secure communication. With the increasing demand for the transmission security in wireless sensor networks, the key management can be done in two methods, providing session key to individual nodes and providing key management to group nodes, before exchanging data securely, encryption keys must be established among sensor nodes. Key usage for secure data transmission, it does not specify how to exchange keys securely. Besides the link layer, upper layers such as the network and application layers also must exchange keys securely. This is a challenging problem because there are many stringent requirements for key management, and the resources available to implement such processes are highly constrained. Many security-critical applications depend on key management processes to operate but also demand a high level of fault tolerance when a node is compromised.

Key distribution refers to the distribution of multiple keys among the sensor nodes. After deployment, each node must broadcast the key's ID number within its communication range to find out the nodes sharing the same key. Therefore, a secure communication can be established as long as there is at least one key being shared. If there is no shared key between two nodes, the link has to be established through two or more key paths.

B. Key Distribution schemes

The [19] three simplest keying models that are used to compare the different relationships between the WSN security and operational requirements are network keying, pairwise keying, and group keying.

Network keying model: It is simple, easy to manage, and uses very little resources. This model allows easy collaboration of nodes; neighbouring nodes can read and interpret each other's data, satisfying the self organization and accessibility requirements. Model has advantage in terms of scalability and flexibility over the other two schemes as there is only one key for the entire network, and it does not change with the addition of nodes. It has unacceptable drawback in robustness exists.

Pairwise keying model: When a new node is added to the network, the node must obtain a new key for communication. Adding new nodes to the network, may affect the flexibility requirement. This is a resource-intensive process that uses much more precious energy when compared with the simple preloading of a network-wide key as in the previous model. Some pairwise key distribution schemes, self-organization comes into question, because they tackle the scalability problem by reducing the number of shared keys, resulting in some nodes being unable to communicate with others and compromising the self-healing and self-organizing abilities of the network.

Group keying: scheme combines the features of both network and pairwise keying schemes. Within a group of nodes that form a cluster, communications are performed using a single, shared key similar to network keying. The communications between group's uses a different key between each pair of groups in a manner identical to the pairwise keying scheme. When one of the nodes is compromised, the compromise of the entire cluster that it belongs to, which is considerably more isolated than the entire network. Scalability is in the form of, increase keys with the number of groups, not with the size of the network. The problem with this scheme is that it is difficult to set up and also the formation of the groups is a very application dependent process. To efficiently distribute the keys, a keying scheme would require group formation information.

V. CLUSTERING SCHEMES

WSN [16] consists of hundreds or thousands of densely populated sensor nodes that sense the data and propagate through the network. They work collaboratively to process and sensed data. These sensor nodes send data streams to base stations either periodically or based on events and base station send the data to the destination node. In a network, sensors nodes may be densely populated wit, the area detected by the sensors are dividing into a number of small clusters. Each cluster has a coordinator or cluster head (CH), and a number of cluster nodes.

A. Base Station

The [18] [19] base station is a powerful node in the wireless sensor network and it can reach a wide range of communication area. The base station can be located at any place of the network, and it is not limited by electric power, memory space, or data-processing capacity. The base station serves as the gateway for external communication. If the base station has been invaded then the whole network will be taken over, so it is assumed that the base station is well protected and can always be trusted.

A sensor node is the core component of a WSN which can take on multiple roles in a network, such as sensing; data storage; routing; and data processing. It is assumed that sensor nodes are randomly distributed, and each node has a unique identity number. Sensor nodes are limited by electric power, memory space, computation capacity, and communication range. Clusters are the organizational unit for WSNs, dense nature of these networks requires the need for them to be broken down into clusters to simplify tasks such a communication.

B. Cluster Head Selection

Each [17] cluster has a coordinator or cluster head (CH). For CH selection any algorithm can be applied. In Cluster formation process, Firstly a cluster head is selected then with the collaboration of BS clusters are formed and finally routing is carried out. The cluster head selection phase starts and all the deployed nodes send their energy levels to the Base Station. Then on the basis of energy level, geographical area and least id cluster head are selected. Network deployment is considered as manual so the base station is well informed about the geographical locations of the nodes. Base Station will select the cluster heads and multicast this information. Cluster heads [18] are the organization leader of a cluster. A cluster [19] head is selected from the sensor nodes in the same cluster, so it has the same capacity and functions as the other nodes. They often are required to organize activities in the cluster. These tasks include but are not limited to data-aggregation and organization the communication schedule of a cluster

C. Cluster communication techniques

There are [17][18] two types communication techniques - a) Intra cluster communication b) Inter Cluster communication. In intra cluster communication, data transmission takes place between the nodes in a same cluster. In Inter Cluster communication, it takes place between the nodes of different cluster.

The clustering technique plays an important role not only for just organization of the network, but also on the network performance. There are several key limitations in WSNs, that clustering schemes must consider.

• Limited Energy: Unlike wired designs, wireless sensor nodes are ”off-grid”, meaning that they have limited energy storage and the efficient use of this energy will be vital in determining the range of suitable applications for these networks. The limited energy in sensor nodes must be considered as proper clustering can reduce the overall energy usage in a network.

• Network Lifetime: The energy limitation on nodes results in a limited network lifetime for nodes in a network. Proper clustering should attempt to reduce the energy usage, and hereby increase network lifetime.

• Limited Abilities: The small physical size and small amount of stored energy in a sensor node limits many of the abilities of nodes in terms of processing and communication abilities. A good clustering algorithm should make use of shared resources within an organizational structure, while taking into account the limitation on individual node abilities.

• Application Dependency: Often a given application will heavily rely on cluster organization. When designing a clustering algorithm, application robustness must be considered as a good clustering algorithm should be able to adapt to a variety of application requirements.

VI. MULTICASTING SCHEMES FOR SECURITY

1. Steiner –based Hierarchical secure multicast Routing protocol

2. Efficient CBMT with mobility aware MDSDV

A. Steiner-based Hierarchical Secure Multicast Routing Protocol

The details of secure multicast routing protocol [23] based on Steiner-based Hierarchical Multicast Routing Protocol, and we introduce the parameters of the proposed communication protocol

1) Nodes information gathering phase: In the secure multicast routing protocol, the source node should verify each node in order to prevent the malicious node to join in the network. We check the authentication of each node by pre-shared key. After the nodes randomly deployed inside, each node sends location information and HMAC to the source node

2) Steiner sub trees distribution phase: For the security reasons, the sensor node should unicast the node-state information table before broadcasting the topology of Steiner sub tree. And the secret keys of source node, Steiner sub tree and cluster need to be included in the message. So the source node executes .

3) Data delivery phase: In the phase of data delivering, the source node transmits multicast packets as the unicast data packet. After the data packet reaches the root node of the sub tree, CHs in the subtree forward the multicast packet in accordance with the order of Height Value. Simultaneously, if the CH detects that it is the destination, it first validate the time stamp T and the value of HMAC in the received message.

Secondly, each CH who is the destination of multicasting broadcasts the content of multicast packet Each MN in the cluster also checks T and value of HMAC. If the inequality holds and computing result is equal to HMAC in the received message, MNsaccept the content of multicast packet.

4) Steiner tree maintains phase: The main task of Steiner tree maintains is re-keying for each node in the network. In our proposed approach, we use the temporary session key to re-key the expired key

B. Efficient CBMT with mobility aware MDSDV

The proposed approach [24] is to achieve secure multicast communication for mobile adoc networks. The Approach uses Multicast version of DSDV routing protocol to maintain routing table periodically. It forms multicast tree among the group members. Each node can determine their present physical location. It quickly adapts to the topology changes. It is used to discover alternate route for failure of existing route. It also sends acknowledgement for each transmission in order to reduce the retransmission. Thus the approach of CBMT using MDSDV tends to have multicast connectivity between the nodes. The approach of Efficient CBMT with mobility aware MDSDV is described in five phases with specific notations.

Phase 1: Authentication: For each node, assign certificate key to verify its node identity. Each node has IP address, node address and certificate key. Certificate key and its IP address encrypt to form a public key. Thus, each node is authenticated based on broadcast request and reply.

Phase 2: Cluster Head Election: Initially the list of Local Controllers (LCs) contains only the source Group Controller GC. Then, GC collects all its 1 hop neighbours by MDSDV routing protocol. Elect LCs which are group members and which have child group members (the LC belongs to the unicast path between the source And the child group members). Verify for each one if it is a group member and if it has child group members then add the LC to the list of LCs. Thus, LCs are selected as cluster heads for its corresponding group members.

Phase 3: Cluster Formation: All the members reachable by this new LC will form a new cluster. If group members that exist and do not belong to the formed clusters then choose the nodes that have the maximum reachability to the others nodes in one hop from the remaining members. This reachability information is collected through the MDSDV routing protocol. Thus, nodes are selected as local controllers for the remaining group members and forms new cluster.

Phase 4: Secure Multicast Communication: The source encrypts multicast data with the TEK, and then sends it to all the members of the group following the multicast tree. The TEK distribution is achieved in parallel, according to the following steps. Initially, the entire group members receive from the source by unicast the session key KEKcsg-0 (key encryption key of the cluster sub-group 0), encrypted with their respective public keys. Each local controller should join this group. The local controllers decrypt this message, extract the TEK, re encrypt it with their respective clusters keys and send it to all their local members.

Phase 5: Node mobility: For frequent node mobility, a new member may join a group or an existing member may leave a group. To ensure secure multicast communication, both forward and backward secrecy has to be maintained. Forward Secrecy: When a node leaves the multicast group, it cannot decrypt the future data. It is known as leave operation. The leave operation is in two cases.

- When an ordinary node leaves, it gives less effect in multicast transmission.
    - When a local controller leaves, it leads to clusterization. It first sends the leave notification to the group controller and then all the members of the current LCs are merged with the other cluster based on the reachability information obtained by the MDSDV routing protocol

C. Comparison between two schemes

Security has the major impact on WSN in Multicast routing ,so two multicast scheme protocol is used to secure multicast data

Protocols	Steiner-based hierarchical secure multicast routing	Multicast version destination sequenced distance vector
Confidentiality	Data is not protected. it can be used by unauthorized	Data is protected .oly authorized users can use it
Fault tolerance	It does not tolerate the fault	It can tolerate the fault
Energy consumption	It consume more energy	less energy

VII. CONCLUSUON

In wireless sensor networks security is the major task to be provided in multicasting So we are using CBMT algorithm based on multicast version of DSDV(destination sequenced distance vector) routing protocol which provides secure communication in tolerating the fault less consumption and reduced packet drop ratio

REFERENCES:

[1]. Cheng-Lung Yang, WernhuarTarng, Kuen-Rong Hsieh and Mingteh Chen. “A Security Mechanism for Clustered Wireless Sensor Networks Based on Elliptic Curve Cryptography” IEEE 2010.

[2]. JaWonKo and Yoon-Hwa Choi “A Grid-Based Distributed Event Detection Scheme for Wireless Sensor Networks Sensors” 2011

[3]. XiaowangGuo ,Jianyong Zhu Research on “Security Issues in Wireless Sensor Networks” International Conference on Electronic & Mechanical Engineering and Information Technology IEEE 2011.

[4]. Ahmad Salehi S., M.A. Razzaque, ParisaNaraei, AliFarrokhtala "Security in Wireless Sensor Networks: Issues and ChallengesUniversity”, Technologies Malaysia Skudai, MalaysiaIEEE 2013.

[5]. Hero ModaresRosliSallehAmirhosseinMoravejosharieh” Overview of Security Issues in Wireless Sensor Networks” Department of Computer system and technology University of Malaya Kuala Lumpur, Malaysia IEEE 2011.

[6]. Abhishek Jain, Kamal Kant M. R. Tripathy “Security Solutions for Wireless Sensor Networks” Department of Computer Science & Engineering ASET, Amity University Noida, India IEEE 2012.

[7]. Bin Tian, Yang Xin Shoushan LU0, Xi ouYang Dong , Li Zhe Gong , Yixian Yang “A Novel KeyMANAGEMENT METHOD FOR WIRELESS SENSOR NETWORKS Information Security Center, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China IEEE 2010.

[8]. Zheng Yue-Feng, Han Jia-Yu, Chen Zhuo-Ran, Li Zheng A novel based-node level Security strategy in wireless sensor network Computer Department of Jilin Normal University BoDa College of Jilin Normal University Si Ping ,China IEEE 2012.

[9]. Adil Bashir, Ajaz Hussain Mir “An Energy Efficient and Dynamic Security Protocol for Wireless SensorNetwork Department of Electronics & Communication Engineering National Institute of Technology, Srinagar, Jammu & Kashmir, India IEEE 2013.

[10]. Thenmozhi, Dr.R.M. Soma sundaram Dean Towards an approach for improved security in Wireless Sensor Networks Department of Sciences SNS College of Engineering Coimbatore, India IEEE 2012.

[11]. SecurityYan-Xiao, Xi’an, Qian-Liang Research On Wireless Sensor Network Telecommunication Engineering Institute Air Force Engineering University Xi’an, Shaanxi, China IEEE 2010

[12]. T.Kavitha, S. JenifaSubhaPriya, Dr.D.Sridharan Design of Deterministic key pre distribution using number theory, Dept of

Electronics & Communication Engg College of Engineering, Guindy, Anna University Chennai, India IEEE 2011.

[13]. JaydipSenSecurity in Wireless Sensor Networks Department of Computer Science & Engineering, National Institute of Science & Technology, INDIA

[14]. David Martins and HervéGuyennetWireless Sensor Network Attacks and Security Mechanisms : A Short Survey Computer Science Department University of Franche-Comté, France IEEE 2010.

[15]. QusayIdreesSarhanaSecurity Attacks and Countermeasures for Wireless Sensor Networks: Survey Department of Computer Science, University of Duhsok, Iraq INPRESSCO 2013.

[16]. S. Jerusha, K.Kulothungan& A. KannanLocation Aware Cluster Based Routing In Wireless Sensor Networks International Journal of Computer & Communication Technology ISSN 2012.

[17]. Ketki Ram Bhakare R. K. Krishna SamikshaBhakareAn Energy-efficient Grid based Clustering Topology for a Wireless Sensor Network International Journal of Computer Applications 2012.

[18]. D. J. Dechene, A. El Jardali, M. Luccini, and A. Sauer A Survey of Clustering Algorithms for Wireless Sensor Networks.

[19]. Cheng-Lung Yang, WernhuarTarng, Kuen-Rong Hsieh and MingtehChenA Security Mechanism for Clustered Wireless Sensor Networks Based on Elliptic Curve Cryptography .IEEE 2010.

[20]. Wang Fangfang, Tao Jun, Shao BiruiAn Energy-Balanced Multicast Routing Algorithm in Wireless Sensor Networks IEEE. 2010

[21] Xin Li, ShuBoQiuResearch on Multicast Routing Protocol in Wireless Sensor Network. IEEE 2011

[22] ShahinMahdizadehAghdam,Mohammad KhansariOn the Better Performance of ADMR versus ODMRP 6'th International Symposium on Telecommunications (IST'2012)

[23] Rong Fan, Jian Chen, Jian-Qing FuLing-Di Ping A Steiner-Based Secure Multicast Routing Protocol for Wireless Sensor Network 2010 IEEE

[24] D.suganya Devi and Dr.Gpadmavathi "Efficient Cluster Based Multicast Communication" International journal Of engineering Science and technology 2010

Secure Multicasting using Geographical Information: A Cluster Based Approach in Wireless Sensor Networks

Manjunath C R[1], Sindhu Anand[2]
[1,2] Department of Computer Science and Engineering
School of Engineering and Technology
Jain University

Abstract:- **In order to guarantee the privacy and safety of data transactions in Wireless Sensor Networks, secure key transportation and unique node identification have become major concerns. WSNs are deployed in a wide range of applications with a high demand for secure communications. When designing a secure key management for a secure communication channel establishment in WSNs, attention should be given to the resource constraints of the devices and the scalability of the network. Here an attempt for public-key to define a hybrid key establishment algorithm for symmetric key cryptography. An Elliptic Curve Cryptography based implicit certificate scheme and how to utilize the certificates for deriving pair-wise link keys in a WSN, by a performance and security analysis. A new approach for secure communication channel establishment is made in order to suite the functional and architectural features of WSNs.**

Keywords: Wireless Sensor Networks, certificate, secure communication, key establishment, Elliptic Curve Cryptography.

I. INTRODUCTION:

Wireless Sensor Networks [3][4][5][14][16] (WSN's) applications are used in various fields from commercial and industrial to military areas. There is a need for security in WSNs, as they communicate happens in an insecure communication medium and they often operate unattended. These devices are economically viable; they have a limited amount of energy, computation power, and memory and communication abilities. A node's lifetime is influenced by the amount of energy that it uses to perform computations and is therefore it's directly influenced by the efficiency of its algorithms. A Pairwise keying process provides basic security services in wireless sensor networks. That enables sensor nodes to communicate securely with each other using cryptographic techniques. Typical public-key cryptography is a low-power domain is for Wireless Sensor Network. The data is transmitted towards the base station using single-hop connectivity comprising of wireless communication links with the nodes, where the data need to be sent in a secure manner. Encryption algorithms for secure transmission are make use of complex algorithm without the key it's impossible to extract information through a cryptanalysis. The classes of cryptographic algorithms can be classified as symmetric or asymmetric. Greater robustness against sensor node does come at a cost, particularly in the overhead involved for key management. If a sensor node communicates with a large number of nodes, it will and must store many keys and select the appropriate ones when communicating. Wireless sensor network in which the nodes are deterministic with similar computational and communication capabilities. The network uses Clustering technique for key distribution and secure communication. In a cluster, all the nodes maintain different keys, but every node uses same key for different communications with the base station.

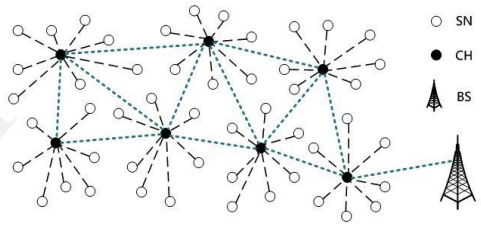


Fig 1: General Architecture of the Network

WSN [15] architecture with cluster is based on different topologies which can be established for heterogeneous and homogeneous networks. The topology in every type of data collecting and monitoring WSNs, there is a cluster head or a network coordinator node. The nodes are deployed in predefined clusters under the control of particular CH. Generally, the CH acts as the intermediate coordinator between sensors and the base station. Sensor nodes gather data from their corresponding area and send data to CH or the base station via single hop or multi-hops. The communication links are very short-term and dynamic changing. For such pair-wise link keys are required for secure communication since symmetric key encryption is more cost effective

II. RELATED WORK / LITERATURE SURVEY:

Key management has remained a challenging issue in WSNs due to the constraints of sensor node resources. Various key management schemes that trade off security and operational requirements have been proposed in recent years.

i. Xiaowang Guo, Jianyong Zhu: [1] has addressed the most suitable asymmetric cryptography primitives for WSN, the Elliptic Curve Cryptography cryptosystem security based on the discrete logarithm problem, and the main ECC primitive operation is the scalar point multiplication.

ii. Ahmad Salehi S, M.A. Razzaque, Parisa Naraei, Ali Farrokhtala: [2] addresses the issues in straight pairwise key participating amongst each two nodes in a large network which have many nodes since. Here an effective key distribution technique in found for preserving the capability to communicate among all related nodes.

iii. Bin Tian, Kamal Kant M. R. Tripathy: [5] addresses about a hierarchical key management scheme to ensure the security of the network services and applications and how it can save the computing and transmitting energy for large-scaled WSNs

III. KEY MANAGEMENT AND KEY DISTRIBUTION:

Key [1] [11] management is one of the most important issues of any secure communication. With the increasing demand for the transmission security in wireless sensor networks, the key management can be done in two methods, providing session key to individual nodes and providing key management to group nodes, before exchanging data securely, encryption keys must be established among sensor nodes. Key usage for secure data transmission, it does not specify how to exchange keys securely. Besides the link layer, upper layers such as the network and application layers also must exchange keys securely. This is a challenging problem because there are many stringent requirements for key management, and the resources available to implement such processes are highly constrained. Many security-critical applications depend on key management processes to operate but also demand a high level of fault tolerance. Multiple Keys are distributed among the sensor nodes; each node must broadcast the key's ID within its communication range to find out if the nodes share the same key. A secure communication can be established as long as there is at least one key being shared. If there is no key shared between two nodes, then the link has to be established through two or more paths.

IV. SECURE TRANSMISSION IN WSN: KEY DISTRIBUTION SCHEMES

There are [11] three keying models that are compared between the WSN security and operational requirements which are network keying, pairwise keying, and group keying. The network keying model has advantages over the pairwise keying, and group keying. It is simple, easy to manage, and uses very little resources. It allows collaboration of nodes where neighbouring nodes can read and interpret each other's data, it is self-organizing, scalability and accessible. The drawback of network keying is robustness.

The pairwise keying it is difficult to add new nodes to the network, hence affecting the flexibility requirement. This is a resource-intensive process that uses more energy when compared with the simple preloading of a network-wide key as in the network keying. Some pairwise key distribution are self-organization, because they tackle the scalability problem by reducing the number of shared keys, results in some nodes being unable to communicate with other nodes and hence compromising the self-healing and self-organizing abilities of the network.

The group keying combines the features of both network and pairwise keying techniques. When group of nodes form a cluster, communications are performed using a single, shared key similar to network keying. The communications between group's uses a different key between each pair of groups in a manner identical to the pairwise keying technique. Scalability, increase keys with the number of groups, not with the size of the network. The drawback with this technique is that it is difficult to set up and also the formation of the groups is a very application dependent.

A. Elliptic Curve Cryptography:

Elliptic curve cryptography [1][8] (ECC) is used for a efficient implementation of a public-key cryptography algorithm, where the security is achieved by using key encryption and decryption to solve the discrete logarithm problem. Most of the Public key cryptography systems are designed based on the RSA algorithm but reaches prefer ECC, because it provides same level of security with much smaller key size and both the public key and private key operation use the same point multiplication operations unlike RSA.

ECC is based on elliptic curves the variables and coefficients which are limited to the elements in a finite field. The finite field is classified by the sizes which are exactly one finite field up to isomorphism of size p^n for each prime p and positive integer n. The elliptic curve in the finite field F_p is a prime curve if p is a prime number, and F_p is defined as a set of integers{0,1,..., p-1}. The elliptic curve is a binary curve in the finite field F_2^m, where m is a large integer andF_2^m is defined as a set of integers {0, 1,..., 2n-1}.

The prime field is Fp. An elliptic curve E is expressed by the equation $y^2 = x^3+ax+b$. The elliptic curve E is defined in the prime field F_p, where the point (x, y) falling into the elliptic curve E will meet the equation

$y^2 = x^3+ax+b(\text{mod } p)$ …… (1)

The above equation is expressed as $E_p(a,b)$, where p is a large prime number and x, y, a, b are the elements of the finite field F_p. Also, a and b must satisfy the following equation

$4a^3 + 27b^2 (\text{mod } p) \neq 0$…….. (2)

B. Diffie Hellman Key Exchange:

The Diffie-Hellman [6][7] key exchange method allows two parties that have no prior knowledge of each other to jointly establish a shared secret key over an insecure communications channel. The key value is used to encrypt subsequent data using a symmetric key cipher. The interpretation of the data packets for verification is performed only when the nodes pass the Diffie Hellman key exchange mechanism, for every non-leaf node is labelled with the its children nodes.

Diffie-Hellman Key Exchange Algorithm

1. Public key, Parameter Creation. A Cluster head chooses and publishes a Prime Public key P and a key c having large prime order in C^*_R.
2. Private key Computations, Node A Choose a secret key a and Compute Node A ≡ c^a (mod P). Node B Choose a secret key b and Compute Node B ≡ c^b (mod P).

3. Public key Exchange of key values
   Node A sends A to Node B → A
   B ← Node B sends B to Node B
4. Further Private key Computations Node A Compute the key value B^a (mod P). Node B Compute the key value A^b (mod P).
5. The shared secret key value is $B^a = (c^b)^a = c^{ab} = (c^a)^b = A^b$ (mod P).

Final key without sharing each other's private random number and c sitting in between will not be able to determine the key as the private numbers were never transferred. The Diffie-Hellman algorithm works perfectly to generate cryptographic keys which are used to encrypt the data being communicated over a public channel.

C. Certificate Generation Method:

A WSN consists of hundreds or thousands of densely populated sensor nodes spread over a medium scaled network, which sense the data and propagate through the network. They work collaboratively to process and sensed data. These sensor nodes send data streams to base stations either periodically or based on events and base station send the data to the destination node. In a network, sensors nodes may be densely populated wit, the area detected by the sensors are dividing into a number of small clusters.

A Grid is a cluster based schemes, in which clusters are equally sized square grids in a two dimensional plane, have a simple structure with less routing management overhead. With the assistance of GPS or localization techniques, the square grid also provides easier coordination among all sensor nodes in the network.

The base station plays a major role in forming the secure transmission channel which acts like a gateway for external communication. The base station gathers information of all the nodes from the Grid to form the clusters according to the location of the nodes. A key management mechanism is used which is based on ECC. It provides authentication services for the identity of nodes and message transmission between the source and destination. It provides a mechanism the add nodes with pre-loaded public keys as certificates to help the other nodes verify their trust worthiness. By doing so, the old nodes do not have to update their keys for secure communications with the new nodes.

Algorithm get Certificate () {
  Every node in the sink node transmission region {
    Request for certificate
    If its valid Authenticated request then
   Certificate Authentication or the Base Station grants certificate to node
    Node gets the certificate}}

Using this algorithm every node in sink transmission range need to ask the permission i.e. certificate from Certificate Authentication to communicate other nodes within its transmission range. The sink node check their authentication if it is valid it grants a certificate to the node otherwise rejected.

Algorithm for send message (){
  If (cluster head sends message to other head) {
    If (sending node is checked for their validation) {
      Message is accepted to receive or route}
  Reject the message}

If any node or header node want to send some information to other header node in the cluster they need to prove their validation based on certificate.

During the node deployment, each sensor node has to go through an initialization phase, where the base station certifies the trust worthiness of the nodes. The base station generates a pair of public and private keys for each node that issues pre-loaded certificates to ensure the trust worthiness of the newly added nodes. Pre-loaded public key can be used as the certificate to ensure the trust worthiness of the newly added nodes to the network. The nodes within a cluster can verify their trust worthiness with each other within the valid period of certificates generation

V. DIFFERENT KEY MANAGEMENT APPROACH:

Key Distribution Approach	Key used
Random	The size of the key ring cannot be small
Deterministic	Graph based stores a key ring. Grid based stores k- dimension key management
Location Based	The location information and direct key is added
Key Agreement Model	Link Compromise Probability
Predistributed keys	approximately linear or quickly increasing to number of compromised nodes
matrices or polynomials	threshold-based
Key Material Deployment Pattern	Local Secure Connectivity
Uniform	Low
Location - based	High

VI. PROPOSED SCHEME:

An attempt is made to provide a secure communication channel for WSN. There are several issues that restrict the network in terms of providing security and managing the network. The key management technique helps to overcome such issues. A grid based network is used the make the network manageable and reduce the computation time while accessing the node. A hybrid cryptography scheme is proposed, where the ECC and DH are combined with a certificate added to the cluster head.

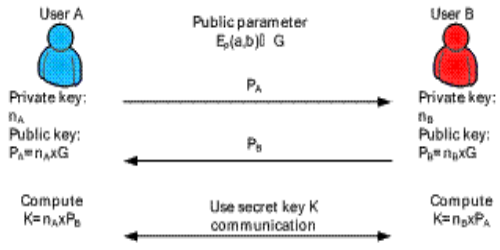


Fig 2: Key Exchange Mechanism.

A. Hybrid Cryptography:

Symmetric [9]key algorithm has a disadvantage of key distribution and asymmetric algorithm need much computation so the power of the sensor is wasted in it and it is not feasible to use as power is wasted then sensor will be of no use. A hybrid cryptosystem is a protocol using multiple ciphers of different types together, each to its best advantage. One common approach is to generate a random

secret key for a symmetric cipher, and then encrypt this key via an asymmetric cipher using the recipient's public key. The message itself is then encrypted using the symmetric cipher and the secret key. Both the encrypted secret key and the encrypted message are then sent to the recipient. The recipient decrypts the secret key first, using his/her own private key, and then uses that key to decrypt the message. The certificate is generated and added to hybrid key management, with ECC+DH. By using this certificate it makes the cluster more secure.

VII. RESULT AND ANALYSIS:

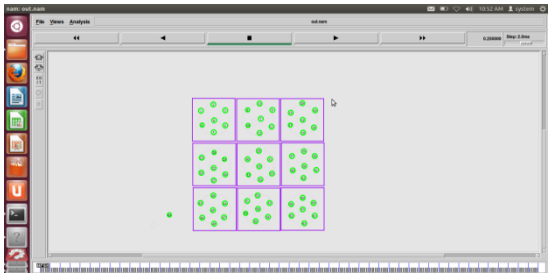


Fig 3: Node deployment.

Node deployment is made deterministic and grids are formed with clusters inside each grid. A cluster head is assigned to each cluster and a Base Station to monitor all the cluster heads as shown in Fig 3.

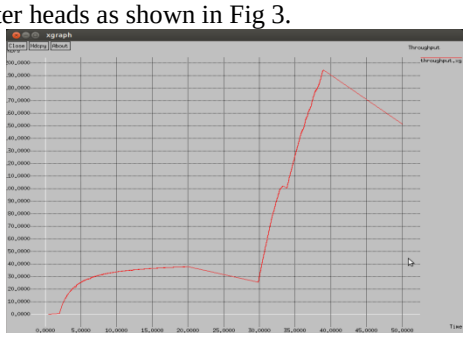


Fig 4: Throughput Graph

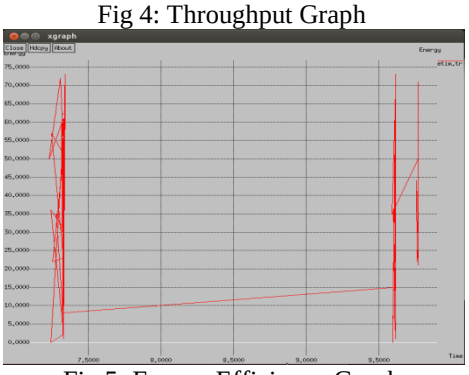


Fig 5: Energy Efficiency Graph

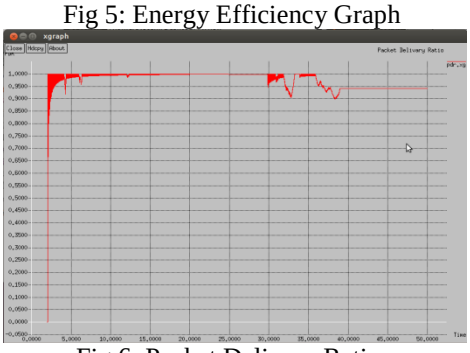


Fig 6: Packet Delivery Ratio.

The proposed system utilizes the energy of the node efficiently, the packet delivery ratio is high and stays moderate due to which the energy consumption is limited. The throughput is high is and hence there is a faster delivery of the data in the network.

VIII. Conclusion:

An attempt as been made to introduced a certificate based pairwise key establishment protocol for WSNs. The proposed key management scheme comprises two phases: For providing certificates for the nodes and establishing pairwise link keys for mutual communication node between the nodes. A secure Public Key Cryptography based solution is derived for a common secret key for symmetric key encryption. The novelty is the utilization of implicit certificates for generating pairwise keys. Our experimental results show the feasibility of deploying the proposed scheme in an actual resource constrained WSN. However, the further optimized ECC operations may have less resource consumptions on sensor nodes and accelerate the protocol execution. Moreover, we have discussed and justified the appropriateness of the protocol for the resource utilization and scalability of WSN. Though there is a simple concept behind the proposed scheme, the security analysis has proven the robustness of the protocol for different security. In future, we intend to extend this protocol by changing the content of the certificate in such way to provide higher security for mobile sensor nodes in massive scale networks. We can customize the content of the implicit certificates by adding other information such as the time stamp, location identity, depending upon the application requirements. The certificates are utilized for group key management in large scale sensor networks.

REFERENCES:

[1]. Cheng-Lung Yang[1], Wernhuar Tarng, Kuen-Rong Hsieh and Mingteh Chen A Security Mechanism for Clustered Wireless Sensor Networks Based on Elliptic Curve Cryptography IEEE 2010.

[2]. Ruan de Clercq, Leif Uhsadel, Anthony Van Herrewege, Ingrid Verbauwhede K.U. Leuven Ultra Low-Power implementation of ECC on the ARM Cortex-M0+, Department of Electrical Engineering and iMinds Kasteelpark Arenberg 10, 3001 Heverlee-Leuven, Leuven, Belgium.

[3]. Gustavo S. Quirino, Admilson R. L. Ribeiro and Edward David Moreno Asymmetric Encryption in Wireless Sensor Networks INTECH 2012.

[4]. Ismail Butun and Ravi Sankar A Brief Survey of Access Control in Wireless Sensor Network IEEE 2010.

[5]. Mohamed Hamdy Eldefrawy1, Muhammad Khurram Khan1, Khaled Alghathbar A Key Agreement Algorithm with Rekeying for Wireless Sensor Networks using Public Key Cryptography. IEEE 2010.

[6]. Geetha, Jayalakshmi Performance Analysis of Sdrp for Wsn Using Diffie – Hellman Algorithm IOSR Journal of Computer Engineering Volume 16, Issue 2, Ver. XII (Mar-Apr. 2014), PP 19-23.

[7]. R.Dhanalakshmi K.Pradeepa Improved Key Selection Techniques for Wireless Sensor Networks IJSR - INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH volume:3 Issue: 4 April 2014.

[8]. Pritam Gajkumar Shah, Xu Huang, Dharmendra Sharma, Analytical study of implementation issues of Elliptical Curve Cryptography for Wireless Sensor networks International Conference on Advanced Information Networking and Applications Workshops IEEE 24th 2010.

[9]. Madhumita Panda Security in Wireless Sensor Networks using Cryptographic American Journal of Engineering Research (AJER) e-ISSN : 2320-0847 p-ISSN : 2320-0936 Volume-03, Issue-01, pp-50-56 2014.

[10]. Srikanta Kumar Sahoo and Manmanth Narayan Sahoo An Elliptic Curve based Hierarchical Cluster Key Management in Wireless Sensor Network 2010.
[11]. Pratik Ranjan Nachiketa Tarasia An Efficient Node Authentication Scheme based on Elliptic Curve Cryptography for Wireless Sensor Networks (IJCSET) ISSN: 2229-3345 Vol. 4 No. 05 May 2013.
[12]. Mr. G. Ravi, Mr. M. Mohamed Surputheen & Dr. R. Srinivasan Fast Energy-Efficient Secure Dynamic Address Routing For Scalable WSNs IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012.
[13]. Xu Huang, Pritam Shah, and Dharmendra Sharma Fast Algorithm in ECC for Wireless Sensor Network VOL II IMECS 2010 March 17-19, 2010.
[14]. Shantala Devi Patil, Vijayakumar B P A Public key distribution and Broadcast Authentication scheme for Wireless Sensor Networks International Conference on Recent Development in Engineering and Technology, 5th August , Mysore,ISBN-978-93-82208-00-6.
[15]. Pawani Porambage, Pardeep Kumar, Corinna Schmitty, Andrei Gurtovz and Mika Ylianttila Certificate-Based Pairwise Key Establishment Protocol for Wireless Sensor Networks 2010.
[16]. A.Ramakrisha , P.VijayaBharathi Secured Dynamic Routing Strategy in Wireless Sensor Networks International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July-2013 179 ISSN 2229-5518.
[17]. Huang Lu, Student Member, Jie Li, Senior Member, Mohsen Guizani, Secure and Efficient Data Transmission for Cluster-based Wireless Sensor Networks, IEEE Transactions On Parallel And Distributed System Year 2013.
[18]. Gicheol Wang, and Gihwan Cho Securing Cluster Formation and Cluster Head Elections in Wireless Sensor Networks International Journal of Communication Networks and Information Security (IJCNIS) Vol. 6, No. 1, April 2014.
[19]. Andrey Khurri, Dmitriy Kuptsov, and Andrei Gurtov On Application of Host Identity Protocol in Wireless Sensor Networks2010

Secure Multi-Keyword Search over Encrypted Cloud Data

Parameshwar Rao D
Department of Computer Science and Engineering,
Jain Global Campus, Jain University, Jakkasandra Post,
Kanakapura Taluk, Ramanagara District, India

S. Balaji
Center for Emerging Technologies,
Jain Global Campus, Jain University, Jakkasandra Post,
Kanakapura Taluk, Ramanagara District, India

***Abstract* — Cloud computing is an emerging technology, providing great flexibility and cost savings. It has motivated data owners to outsource their data from local systems to commercial public cloud. But sensitive data has to be encrypted before outsourcing to cloud to secure such data. This requires the traditional method data utilization based on plain text keyword search to be augmented with additional feature of searching encrypted data. Considering the large number of data users and data in cloud, it is necessary for the search service to allow multi-keyword query and return the data in the order of their relevance to these keywords. However, retrieving all the files having queried keyword will be expensive in "pay as you use" cloud paradigm. In the proposed system, vector space model and homomorphic encryption are employed wherein, the vector space model helps to provide sufficient search accuracy and the homomorphic encryption enables cloud service providers to perform the search operation based on user's multi-keyword search query without the need to decrypt it and also enables users to involve in the ranking.**

Keywords — Cloud, Ranking, Homomorphic Encryption, Vector Space Model.

1. INTRODUCTION

Cloud computing is a promising service for data outsourcing and high quality data services. Its great flexibility and cost savings are motivating both the individual users and the enterprises to outsource their local complex data management system into the cloud. To protect data privacy and combat unsolicited accesses in the cloud, sensitive data such as e-mails, personal health records, photo albums, tax documents, financial transactions, etc. may have to be encrypted by data owners before outsourcing to the commercial public cloud. However, this obsoletes the traditional data utilization service based on plain text keyword search. Thus, in order to ensure privacy of the personal information over the cloud, data owner must encrypt the data before uploading it to the cloud. There is a problem faced by the users since the cloud service provider needs to perform the calculations on data in order to respond to the requests made by the user. User has to provide key to the cloud service provider to decrypt the data before executing the required calculations, this affects the confidentiality of data stored in the cloud by the data owner.

In cloud computing, data owner may share their outsourced data with a number of users who might want to retrieve the data files of their interest. One of the most popular ways to do so is through keyword-based retrieval. Keyword-based retrieval is a typical data service and widely applied in plain text scenarios, in which users retrieve relevant files in a file set based on keywords. This searchable encryption schemes are impractical for real world cloud computing scenarios because these systems are designed to handle either a single keyword search or a Boolean search. The main drawback of single keyword-based search is that, it is hard to express complex information needs. On the other hand, Boolean search uses the presence or absence of queried keywords to retrieve the matched documents [2] and it is very difficult for most of the users to control the number of retrieved documents. In contrast, Information Retrieval (IR) systems [8] utilize ranked-search model to rank all the retrieved documents according to some relevance criteria. Such a model provides a precise answer by retrieving only the top−k relevant documents from the whole document collection. Furthermore, for efficiency purposes, all the current searchable encryption schemes reveal the access pattern to the un-trusted cloud service providers.

In order to improve feasibility and save the cost, it is preferred to get the retrieval result with the most relevant files that match user's interest instead of all the files. This indicates that the files should be ranked in the order of relevance by user's interest and only the files with the highest relevance need to be sent back to the users.

Currently, files are ranked only by the number of retrieved keywords. This impairs search accuracy and security is implicitly compromised to tradeoff for efficiency, which is particularly undesirable in security-oriented applications. Preventing the cloud from involving in ranking and entrusting all the work to the user is a natural way to avoid information leakage. However, the limited computational power on the user side and the high communication overhead precludes information security [5].

The homomorphic encryption enables users to involve in the ranking while the majority of computing work is done on the server side by performing search operation only on cipher text. Along with sending the encrypted data

over cloud, the data owner may also send the searchable index. Searchable index is a collection of phrases and keywords, which facilitates fast and accurate information retrieval. Thus, storage of searchable index along with the encrypted data in the cloud optimizes speed and performance in finding relevant documents for a search query.

In the proposed approach, searchable index is built from the collection of files that needs to be stored over the cloud in order to facilitate the fast and accurate retrieval of data. Homomorphic encryption and vector space model that guarantees the retrieval of most relevant data by performing user's multi- keyword search operation over encrypted cloud data are employed.

2. RELATED WORK

Ning Cao and Cong Wang [1], establish a set of strict privacy requirements for a secure cloud data utilization system. Among various multi-keyword semantics, authors use the efficient similarity measure of "coordinate matching", i.e., as many matches as possible, to capture the relevance of data documents to the search query and further use "inner product similarity" to quantitatively evaluate such similarity measure. The drawback observed is that direct outsourcing the data vector or the query vector will violate the index privacy or the search privacy.

Peng lu, Jiadi Yu, Xin Dong [2], introduce Two Round Searchable Encryption (TRSE) which preserves privacy of data retrieved but at higher communication overhead which has direct impact on efficiency.

Ayad Ibrahim, Hai Jin, Ali A. Yassin, Deqing Zou [3], propose a scheme that uses two distinct cloud servers, one for storing the secure index, while the other is used to store the encrypted document collection. Such a new setting prevents leaking the search result, i.e. the document identifiers, to the adversary cloud servers. The drawback is that utilizing two cloud servers is expensive which makes it impractical to use.

Bharath K, Samanthula and Wei Jiang [4], propose an efficient method for converting an encrypted integer z into encryptions of the individual bits of z and security primitive to construct a new protocol for secure evaluation of range queries in the cloud computing environment. Also, authors employ Privacy-Preserving Range Query (PPRQ) protocol which protects the confidentiality of the data and input query but reveals data access patterns.

Jiadi Yu, Peng Lu, Yanmin Zhu and Guangtao Xue [5], formulates the privacy issue from the viewpoint of similarity relevance and scheme robustness. It is observed that server-side ranking based on Order-Preserving Encryption (OPE) inevitably leaks data privacy. Data updates like adding or deleting files lead to a new challenge to the searchable encryption scheme.

Maha Tebaa, Said El Hajji, Abdellatif El Ghazi [6], propose a method to perform the operation on encrypted data without decrypting it and show that the same result as well when the calculations were carried out on the raw data but the efficiency is a tradeoff.

3. PROPOSED SYSTEM

The proposed approach retrieves the data using multi-keyword search over encrypted cloud data. Ranking is left to the user side in order to achieve data privacy.

In this approach, data owner uploads both encrypted files and the searchable index to the cloud server. As shown in Figure 1, data user sends a query consisting of multi-keywords to the cloud service provider. These queries will be processed by the cloud service provider by computing the scores from the encrypted index stored on the cloud and then cloud service provider will return the encrypted scores of files to the user. Next, the user decrypts the scores and picks up the top-k highest scoring file identifiers and request to the cloud server. Finally, the user gets the search result from the cloud server.

In order to reduce the computational burden on the user side, computing work is done at the cloud service side. It is essential to choose an encryption scheme that guarantees the operability and security on the server side.

Homomorphic encryption allows only specific types of computations to be carried out on the corresponding cipher text. Though the original fully homomorphic encryption property has such a fine property, which employs ideal lattices over a polynomial ring [8] it is complicated and inefficient for practical utilization.

In the fully Homomorphic Encryption Over the Integers (FHEI) scheme [7], the approximate integer Greatest Common Divisor (GCD) is incorporated to provide sufficient security. The cipher text resulted from the encryption of data will be large in size. In order to reduce the size of cipher text and the communication overhead, the original FHEI scheme should be modified to be more flexible in order to ensure the correctness of the decryption. Fortunately, as a result of employing the vector space model in top-k retrieval, only addition and multiplication operations over integers are necessary to compute the relevance scores from the encrypted searchable index. In the proposed approach, the original homomorphism in a full form is reduced to a simplified form that only supports integer operations, which achieves better efficiency than the full form. To further reduce the communication overhead, an ElGamal encryption is employed that result in generation of cipher text with smaller size.

Figure 1: Retrieval of Encrypted Cloud Data

Thus, the proposed system ensures the secure multi-keyword search over encrypted cloud data with improved efficiency.

4. ALGORITHMS USED

The algorithms that are used in the proposed system are as follows:

Algorithm 1: ElGamal

It is an asymmetric algorithm. That is, it makes use of two keys- (i) public key to encrypt the files and (ii) secret key to decrypt the files.

The steps followed in this algorithm are as follows:

ElGamal Key Generation Protocol:

The key generation protocol for the ElGamal algorithm is as follows:

Step 1: Start
Step 2: Create a random prime number, p. This number is the ElGamal "modulus".
Step 3: Select two other random numbers, g and x, both of which are less than p; these numbers do not have to be prime.
Step 4: Compute y, where $y = g^x \bmod p$. The public key is p, g and y. The private key is x.
Step 5: Stop

ElGamal Encryption Function:

The ElGamal encryption function is as follows:

Step 1: Start
Step 2: Break the plaintext into blocks based on the length of the public key modulus.
Step 3: Process each plaintext block, m:
a. Choose a random number, k, relatively prime to p - 1 (i.e., k and p - 1 have no common factors).
b. Compute $a = g^k \bmod p$.
c. Compute $b = (y^k m) \bmod p$.
d. Concatenate a and b to form an encrypted data block.
e. Concatenate the encrypted data blocks to form the cipher text.
Step 4: Stop

ElGamal Decryption Function:

The ElGamal decryption function is as follows:

Step 1: Start
Step 2: Break the cipher text into blocks that are twice the length of the key modulus
Step 3: Process each cipher text block:
a. Break the cipher text block in half to form a and b
b. Use the private key to compute the decrypted block m where

$$m = \frac{b}{a^x} \bmod p$$

c. Concatenate the decrypted blocks to form the plaintext.
Step 4: Stop

Algorithm 2: TopkSelect (source, k)

This algorithm is used to retrieve only the top-k ranked file list from the cloud server which is the result of the search operation according to the data user's multi-keyword search query.

The steps followed in this algorithm are as follows:
Step 1: Start
Step 2: Set topk = 0; topkid = 0;
Step3: Begin loop for all item ∈ source do
Step 4: INSERT (topk, (item, itemindex))
Step 5: end for loop.
Step 6: Begin loop for all tuple ∈ topk do
Step 7: topkid.append(tuple[1])
Step 8: end for loop
Step 9: return topkid
Step 10: Stop

Algorithm 3: Insert (topk, (item, itemindex))

This algorithm is used to insert/store the keywords, in order to build a searchable index. Searchable index is a collection of keywords that facilitates fast and accurate retrieval of data.

The steps followed in this algorithm are as follows:
Step 1: Start
Step 2: Condition check if length (topk) < k then
Insert (item, item index) into topk in non-decreasing order of item
Else if condition fails then continue
Step 3: Bein loop for all element ∈ topk do
Step 4: if item < element [0] then
Continue
Step 5: else if condition fails then
Step 6: Discard topk [0], insert (item, item index) into topk in non decreasing order of item
Step 7: end if condition
Step 8: end for loop
Step 9: end if condition
Step 10: Stop

Algorithm 4: Porter Stemmer

Porter Stemmer is one of the algorithms that are used in the information retrieval to reduce the size index files. A single stem typically corresponds to several full terms by storing stems instead of terms compression factors are achieved.

The steps followed in this algorithm are as follows:
Step 1: Start
Step 2: Gets rid of plurals and –ed or –ing suffixes.
Step 3: Turns terminal y to i when there is another vowel in the stem.
Step 4: Maps double suffixes to single ones: –ization, -ational, etc.
Step 5: Deals with suffixes –full, -ness, etc.
Step 6: Takes off –ant, -ence, etc.
Step 7: Removes a final –e.
Step 8: Stop.

5. Security Analysis

The foremost thing to be analyzed in the proposed system is that cloud server should not be able to know the

content, either of the data files, searchable index or the search keyword queries. Secondly, the cloud server should not be able to know the similarity relevance of terms or files so that the proposed system is highly robust.

The proposed system is able to conceal the access pattern and search pattern to be hidden from the cloud server; that is, if suppose the same keyword "t" is requested in two different queries as REQ1 and REQ2. Then, it forms the corresponding query vector say T1 and T2. After that, REQ1 and REQ2 are encrypted into two different cipher texts. Thus, same keywords in different queries are independent to each other, which mean that the keywords retrieved are hidden; thus, the access pattern and search pattern are secure.

In the proposed approach, an ElGamal encryption algorithm is used, it reduces the size of cipher text resulted by encrypting the plain text. As the size of cipher text is reduced, the communication overhead between the cloud service provider and the data user will also be reduced which improves the efficiency.

6. CONCLUSION

The proposed work ensures the secure multi-keyword search and top-k data retrieval over encrypted cloud data. The majority of computing work is carried out by server side by performing operations on cipher text which reduces computation overhead on data user side.

REFERENCES

[1] Ning Cao, Cong Wang, Ming Li, Kui Ren and Wenjing Lou, "Privacy-Preserving Multi-keyword Ranked Search over Encrypted Cloud Data", IEEE: 2011

[2] Peng Lu; Jiadi Yu; Xin Dong; Guangtao Xue; Minglu Li "Privacy-Aware Multi-Keyword Top-k Search over Untrust Data Cloud", 18th International Conference on Parallel and Distributed Systems (ICPADS), pp.252 – 259, 2012

[3] Ayad Ibrahim, Hai Jin, Ali A. Yassin, Deqing Zou, "Secure Rank-ordered Search of Multi-keyword Trapdoor over Encrypted Cloud Data", published at Asia-Pacific Services Computing Conference (APSCC), 2012

[4] Bharath K, Samanthula and Wei Jiang, "Efficient Privacy-Preserving Range Queries over Encrypted Data in Cloud Computing", IEEE: 2013

[5] Jiadi Yu, Peng Lu, Yanmin Zhu and Guangtao Xue, "Toward Secure Multi keyword Top-k Retrieval over Encrypted Cloud Data" IEEE: 2013

[6] Maha Tebaa, Said El Hajji, Abdellatif El Ghazi, "Homomorphic Encryption method applied to Cloud Computing", IEEE: 2012

[7] M. Van Dijk, C. Gentry, S. Halevi, and V. Vaikuntanathan, "Fully Homomorphic Encryption over the Integers," Proc. 29th Ann. International Conference, Theory and Applications of Cryptographic Techniques, H. Gilbert, pp. 24-43, 2010

[8] C. Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," Proc. 41st Ann. ACM Symp, Theory of computing (STOC), pp. 169-178,2009

Securing Data Against Botclouds and IP Spoofing

Usha L[1], Chidananda Murthy P[2]
[1]M.Tech, Dept. of CSE, SET, Jain University
[2]Assistant Professor, Dept. of CSE, SET, Jain University

***ABSTRACT*- Cloud computing is the new trend in computing and resource management. Cloud offers great benefits in terms of flexibility, scalability and availability. Few people and organizations step back from adapting cloud technology mainly because of the security concerns. The security breaches of cloud computing include DoS attacks. BotClouds and IP Spoofing are two DoS attacks. Traditional network security techniques cannot be used to solve cloud computing threats because of the huge and variety of data in cloud. IP spoofing is presenting a false truth in a credible way to gain unauthorized access to cloud services. BotClouds are cloud based botnets which is the most commonly used platform attackers use to perform frauds in cloud environment. This paper proposes a methodology to secure cloud services by monitoring the traffic and logging the activities even for short periods.**

Keywords: IP Spoofing, BotClouds, DoS attack, Cloud computing

I. INTRODUCTION

Moving data to a Cloud environment presents an opportunity to achieve tremendous cost savings compared to the cost to purchase an equivalent amount of data for a locally hosted data center. As with virtual machines, a customer's data is stored over a shared infrastructure that may be distributed throughout multiple Cloud data centers. Adequate security measures must be in place to ensure unauthorized users cannot access data either intentionally or accidently. One of the potential draw backs of moving data processing to a Cloud environment is losing direct control Monitoring at the OS or VM level is a basic means to monitor your systems but to closely monitor attacks tools such as an Intrusion Detection System (IDS) should be used. It is essential for the cloud service providers to maintain the data security to its clients for cloud computing and internet to reach their full potential. Inefficient data security measures will reflect on the cloud performance and reputation of the cloud provider. Clients expect their data and applications stored in cloud to remain private and secure. As the challenges of security and privacy are evolving along with cloud, security is responsibility of both the customer and the service provider.

IP spoofing is a technique used to gain unauthorized access to cloud services where by the attacker sends a request with a forging IP address [10] indicating that the request is coming from a trusted host. Many cloud applications use the IP address of a computer to identify a valid entity. In certain cases, it is possible for an IP address to be falsely assumed Attacker puts an internal or trusted IP address as its source IP address [13]. The cloud sees the IP address as trusted and lets it through. A hacker uses a valid IP addresses.

BotClouds are cloud based botnets [1]. Bots are collection of scripts that perform automated tasks in cloud. Rather than use a network of infected machines, Bots use Cloud services to build BotCloud.Botmasters registers to the CSP and introduce bots to the applications and data hosted on cloud. Bots cannot be noticed [5] easily because they perform tasks as similar to humans but at a higher rate than humans. Hence it is difficult to detect bots using an intrusion detection system (IDS), so it is appropriate to detect bots based on their behavior [4]. Bots are frequently used to launch DoS attacks [6].

II. RELATED WORK

Surveys from Gartner [9] have shown that security is the main obstacle in cloud computing [3]because of which organizations do not adapt to cloud computing. Main reason for this is cloud computing facilitates the storage of data at remote site to maximize resource utilization. Bots act safe and are hidden as long as possible and they are easy to establish when compared to a traditional botnets. Firewalls are inefficient in preventing and detecting bots. Hence active monitoring of network traffic of anomalous activity is suitable to detect bots. A DoS attack can be launched by a BotCloud. Botnets bring down relatively unprotected websites just be directing thousands of traffic requests. DoS attack is accompanied by IP Spoofing so as to hide the source of flooding and to make every request look different. Source IP spoofing attacks are critical issues to the Internet. These attacks are considered to be sent from bot infected hosts.The cloud security alliance has initiated research group called Anti-Bot Working group to detect bot attacks on cloud using new and efficient methodologies. Cloud Security Alliance has also

mentioned threats to Cloud Computing [11] which includes BotClouds and IP Spoofing attacks.

Traditional security technologies lack the sophisticated capabilities and visibility required to detect and protect cloud data from attacks in real time monitoring of cloud, there is a need to analyze who is accessing which data from which resource at what time Or do we have a breach of compliance standard C because of action A? Traditional network security techniques [12] for detecting Bots and IP spoofing are inadequate to solve cloud attacks because

- Traditional tools were not designed to analyse and manage huge amount of data.
- Though traditional network security systems collect logs and events from a huge variety of systems they cover only a part of potentially relevant activity.
- Detection of threat relies on having signatures or knowing methods of attack in advance.

CAPTCHA- Computer Automated Public Turing test to tell Computers or Human Apart are used to differentiate between human and bots. They are used to prevent bots but recent surveys have reported that Captchas can be easily cracked by using advanced character and pattern recognition software's [14]. Websense security Labs have reported that Windows Live Captchas can be cracked in as little as 60seconds. Captchas are not a good choice because they are difficult for blind andpartially sighted people and bad Captchas are easy to decode where as good Captchas are difficult to decode.

III. PROPOSED WORK

The proposed architecturesintend to detect requests from Spoofed IP addresses and detect Bot activity by continuous behavior based monitoring.

A. SPOOF DETECTION

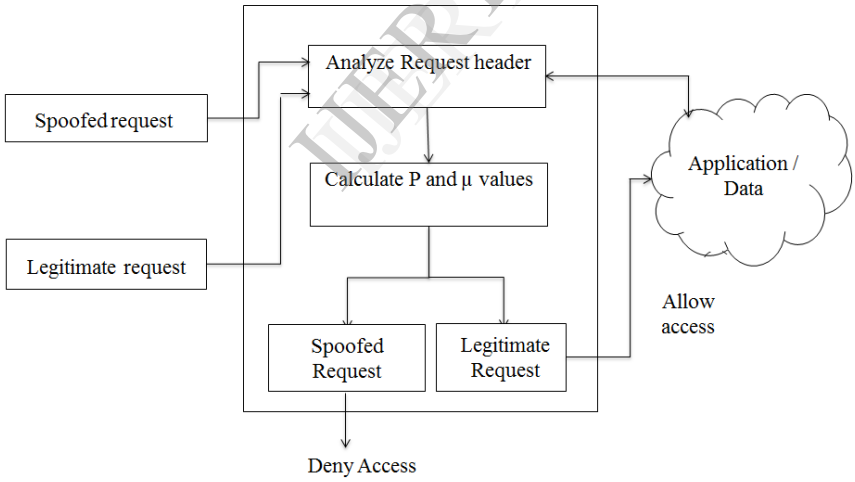


Fig.1 Proposed architecture to detect IP Spoofing

Spoofing the IP in cloud is done by customizing the HTTP header by manipulating the original source address IP with the forged IP address while sending the request to access the cloud data. The X-Forwarded-For is a HTTP header field to identify the originating IP address of a client connecting to a web server. This field can be forged to spoof the IP address and send requests. The X-Request-Start is a filed in the http header which holds timestamp at which the request was created.

The spoof detection module examines the HTTP header and detects the IP address in the header. It also examines for time the request was created and the time at which the request has reached the destination. According to the ANT algorithm to detect IP spoofing, legitimate packets choose the shortest route and reach the destination in less time where a spoofed packet does not take the shortest route and hence takes more time to reach the destination.Pheromone is the increase in possibility of choosing a path based on the number of requests previously chose the path

The X-Request-End time, X-Request-Start time, and the ping time is calculated for each arriving request. The difference of the X-Request-End and X-Request-Start is the time taken by the request to reach the destination. Then the Ping time and time taken is calculated as Difference in time. Initially the Pheromone is set to 0. If the Pheromone value (P) is 0, the difference in time value obtained is set as the pheromone value else the average of existing Pheromone value and Difference in time is set as the updated Pheromone value.The requests are categorized as Spoofed (µs) or legitimate (µleg). The request from Spoofed IP address is denied and the request access from legitimate IP address is granted.

B. BOT DETECTION

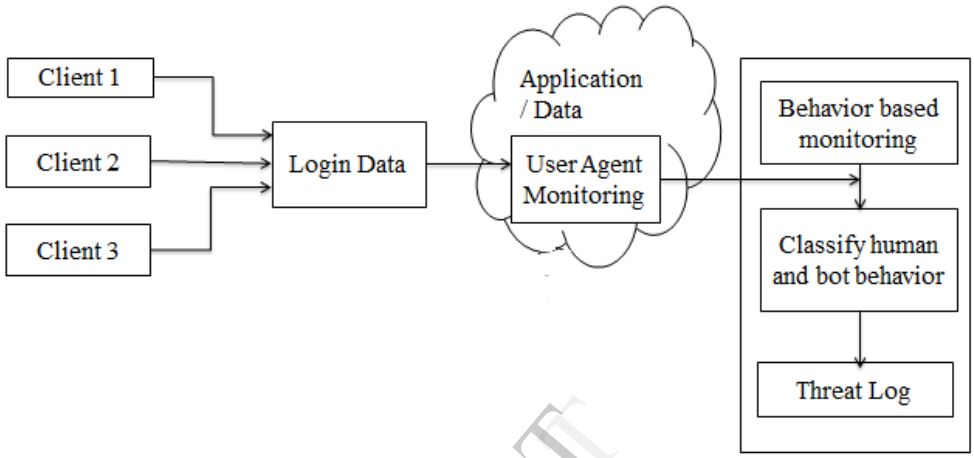


Fig.2 Proposed architecture to detect Bot activity in cloud

A traditional botnet requires substantial time to build, whereas a BotCloud can be online in minutes. The bots enter the cloud from any of the client machine, hence continuous monitoring of the network traffic is necessary to detect threats in cloud [2]. Bots attack the application by performing fraud activities at a very faster rate than humans. It is difficult to detect bots hence the details of the user agents like operating systems, browsers are also stored. The network traffic generated by activities on the cloud are monitored based on action time and action frequency[7] for parameters like number of clicks, file requests, failed requests html to image ratio and many more. The activities which generate abnormal network traffic and the user agents logged in are joined together to find out the IP address of the client from which bots are entering the cloud. The log of IP addresses is maintained and if the administrator wishes, an information email can be sent to the user logged in from the infected IP address.

IV. RESULTS

The proposed methodology, which monitors the traffic generated by applications, hosted on cloud continuously, rather than monitoring the past traffic patterns are better. The Spoofed IP is detected by examining the HTTP header and categorization of spoofed and legitimate request is done based on time factor i.e. the time taken for the request to reach the destination. The administrator has an option ofinforming the clients about the suspicious activity detected from the clients IP address through email.

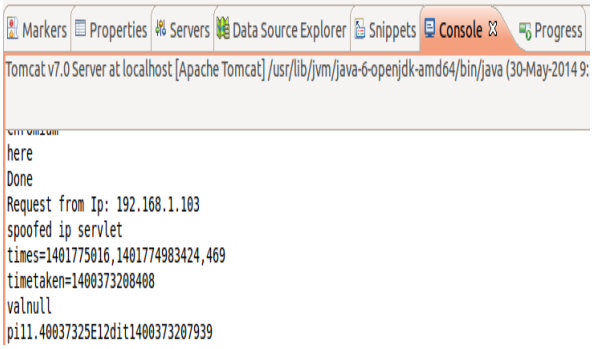


Fig.3 Detection of Spoofed IP

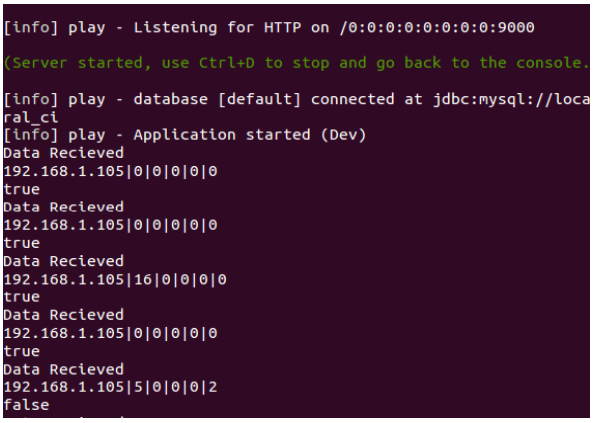


Fig.4 Continuous traffic monitoring for bot activity

V. CONCLUSION AND FUTURE ENHANCEMENT

The proposed methodology detects, request from spoofed IP address and the presence of bots in the cloud. The user agent's properties are stored to track the client's. Currently monitored bot activities are click frauds and frequent inappropriate data submitted in the form. The IP address from which bots are entering are detected and logged in the database. The administrator can view the infected IP address and sends an information email to the user regarding malicious activity. The existing work can be extended to monitor more complex bot activity such as ban enforcement activity which finds and reverts changes, automatic importer by request for all types of document formats which has the same content.

REFERENCES

[1] Kassidy Clark et al., "BOT-CLOUDS - The Future of Cloud-based Botnets", 2010.
[2] Mr. D. Kishore Kumar et al., "Cloud Computing: An Analysis of Its Challenges & Security", in IJCSN, 2012.
[3]Sateesh Kumar Peddojuet.al.,"Packet Monitoring Approach to Prevent DDoS Attack in Cloud Computing", in IJCSEE, 2012.
[4] Shun-Wen Hsiao, Yi-Ning Chen, "A Cooperative Botnet Profiling and Detection in Virtualized Environment", in IEEE, 2013.
[5] Keith Harrison1, BehzadBordbar, "A framework for detecting malware in Cloud by identifying symptoms", IEEE, 2012.
[6] MatijaStevanovic Jens Myrup Pedersen, "Machine Learning Based Botnet Detection", 2013.
[7] PedramHayatiet.al, "Behavior-Based Web Spambot Detection by Utilizing Action Time and Action Frequency", Springer, 2010.
[8] Kuochen Wang,et al., "A fuzzy pattern-based filtering algorithm for botnet detection, Elsevier journal", 2011.
[9] NarayananArumugam and Venkatesh, "Triangular fuzzy based classification of IP request to detect spoofing request in data network" academic journal of International Journal of Physical Sciences, 2013
[10].http://www.proofpoint.com/solutions/threat-management.php
[11].http://www.mediabuzz.com.sg/asian-emarketing/march-2011/1230-poofing-attacks
[12]https://cloudsecurityalliance.org/research/big-data/
[13]http://www.bankinfosecurity.com/webinars/unknown-threats
[14]https://blogs.oracle.com/vreality/entry/public_cloud_security_anti_spoofing2
[15]http://arstechnica.com/security/2008/04/gone-in-60-seconds-spambot-cracks-livehotmail-captcha/

Sharing Secure Data in The Cloud To Multiple users Among a Dynamic Group

Mahesh A.
Department of Computer Science and Engineering
Jain Global Campus, Jain University, Jakkasandra Post
Kanakapura Taluk, Ramanagara District-562112

S. Balaji
Centre for Emerging Technologies
Jain Global Campus, Jain University, Jakkasandra Post
Kanakapura Taluk, Ramanagara District-562112

Abstract: **Cloud computing provides an economical and efficient solution for sharing group resources among cloud users. When sharing the data in a group while preserving data, identity privacy is still a challenge because of frequent changes in the membership. To overcome this problem, a sharing secure data scheme among multiple users among a dynamic group is proposed so that any user within a group can share the data in a secure manner by leveraging both the group signature and dynamic broadcast encryption techniques. It enables authorized users to anonymously share data with others within the group. It supports efficient member revocation and new member joining the group. User revocation list is performed by group manager and it is given to the cloud service provider to check the active users within the group before giving access to the cloud.**

Keywords: Cloud, Data Sharing, Dynamic Groups, Revocation List

1. INTRODUCTION

Cloud computing is recognized as an alternative to traditional information technology due to its intrinsic resource-sharing and characteristics of low-maintenance. In cloud computing, many computing resources are provided as services over the Internet. One of the main services provided by cloud is storage, which allows users to upload and store their data in the cloud. Storing data in cloud provides many advantages such as reliability and availability, but it also brings many other challenges such as secure data sharing.

An organization allows its group members to store and share their data files by utilizing the cloud. Group members can be completely relieved from local data storage and maintenance, but significant risk arises in confidentiality of those stored files. The cloud users are not fully trusted since the cloud servers are operated by cloud service providers. Confidentiality of the data is very important because of the sensitive data stored in the cloud. To preserve the data confidentiality, a basic solution is to encrypt data files and then upload the encrypted data into the cloud.

Firstly, identity privacy is one of the major issues in cloud. Without the identity privacy, users are not willing to join the cloud because their real identities could be easily disclosed to cloud providers and attackers.

Secondly, it is highly recommended that any member in a group should be able to fully enjoy the data storing and sharing services provided by the cloud, which is defined in the multiple-owner manner. Compared with the single-owner manner, where only the group manager can store and modify data, in multiple-owner manner each user in the group is able to not only read the data but can also modify the part of data in the entire data file shared by the users.

Thirdly, groups are normally dynamic in practice. It does not support new user participation and current employee revocation within the group. So the complexities of user participation and revocation in these schemes are linearly increasing with the number of data owners and the number of revoked users.

Lastly, group owner must be able to trace if any disputes occurs within the group.

2. RELATED WORK

Junod and Karlov [1], propose a "CP-ABE based broadcast encryption" scheme that supports direct user revocation. In this scheme, each broadcast receiver's identity is mapped to an individual attribute. The access policy consists of a set of system attributes with a set of identity attributes. Individual user revocation is achieved by updating the set of identity attributes in the access policy.

B. Wang et. al. [4], focuses on "cloud computing and storage services". Accordingly, cloud data is not only stored in the server, but routinely shared among a large number of users in a group. In this paper, the authors propose Knox, a privacy-preserving auditing mechanism for data stored in the cloud and shared among a large number of users in a group.

S. Yu, C. Wang, K. Ren, and W. Lou [6] present a "scalable and fine-grained data access control" scheme in cloud computing based on the Key Policy Attribute Based Encryption (KP-ABE) technique. In this scheme, the data owner uses a random key to encrypt a file, where the random key is further encrypted with a set of attributes using KP-ABE. Then, the group manager assigns an access structure and the corresponding secret key to authorized users such that a user can decrypt a cipher-text if and only if the data file attributes satisfy the access structure

Kamara et. al. [10] propose a framework of a "Cryptographic Storage Service (ACSS)" which considers

the issue of building a secure cloud storage service on cloud infrastructure where the service provider is not fully trusted by the user. It is made up of three basic components (DP, DV, TG) and realizes encryption storage and integrity validation by a group of protocols. However, ACSS is hard to build since it deals at a high level and requires modification of large amount of source code of cloud storage platform.

3. PROPOSED SCHEME

The proposed scheme is to secure the data against unauthorized access by enforcing access control mechanisms. To achieve secure data sharing for dynamic groups in the cloud we combine both the group signature and dynamic broadcast encryption techniques. The short group signature introduced by Chaum and van Heist is used, which enables any users in a group to anonymously use the cloud resources provided by cloud service provider. It supports efficient user revocation and provides secure and privacy-preserving access control to users that guarantee any member in a group to anonymously utilize the cloud resource. Moreover, the real identities of data owners can be revealed by the group manager when disputes occur within the group using the group signature. Group manager will have the list of active users and also maintains the list of revoked users. Only the members of the group can create valid group signatures. Figure1 shows how the group members register with the group owner and how the data is shared among the group members from the cloud server.

The dynamic broadcast encryption is another technique which allows data owners to securely share their data files with other users in the group including newly joined users. It supports new member joining the group without updating private keys of remaining users in the group.

The Dynamic broadcast encryption allows broadcaster to distribute the data only to set of users who requested the data and each user has to compute revocation parameters to protect the confidentiality of the data from the revoked users. The dynamic broadcast encryption scheme is used such that revoked users cannot access the data once they are revoked from the group. Group manager can enable the revoked users to rejoin the group again. The group manager is allowed to compute the revocation parameters, which includes the list of revoked users and make this revocation list available to public by migrating them into the cloud. Each time when users request for data, cloud service provider verifies the revocation list and then provide access to data only to active users in the group. Such a design significantly reduces the computation overhead.

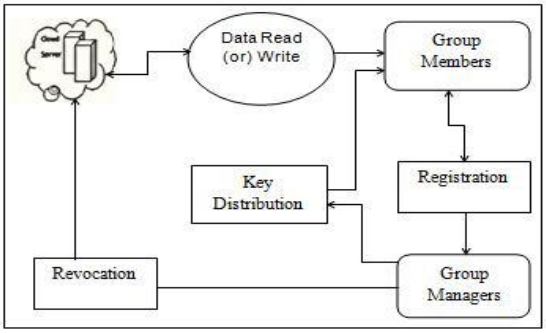


Figure 1: System Architecture

Group manager takes charge of system initialization, signature generation, user registration, user revocation, and revealing the real identity of a user when the dispute occurs.

Data files are encrypted using key policy attribute based encryption. Private and public keys are generated by the users. Public key for encrypting the data and private key is kept secret and is used for decrypting the data files.

The proposed scheme also supports new member joining the group and user revocation from the group. User revocation is performed by the group manager via a publicly available Revocation List (RL) without updating the private keys of remaining users. Each user must follow the revocation parameters before revoking from the group.

4. ALGORITHMS USED

The algorithms used in the proposed system are as follows:

Algorithm 1: Signature Generation

This algorithm is used to generate the signature for the members in the group, through this group signature users are allowed to login into the group.

Step1: start
Step2: Input: Private Key (Ai, xi), system parameter (P, U, V, H, W) and data M.
Step3: Output: Generate a valid group signature on M.
Step4: begin
Step5: Select random numbers Set (t1, t2, t3, r1, r2, r3)
And set x1=a and x2= b
Step6: Compute the following values t, t2, t3, r1, r2, r3.
Step7: compute the challenging c
c<- h (m, t1, t2, t3, r1, r2, r3, r4, r5) using Hash function.
Step8: using c construct the Values s, s2, s3, s4, s5
Step9: Output the signature computed as gsp<-(t1,t2,t3,c,s1,s2,s3,s4,s5)
Step10: stop.

Algorithm 2: Signature Verification

Algorithm 2 is used to verify the group sign and individual user sign during the data sharing from the cloud server.

Step1: start
Step2: Input: System Parameter (P, U, V, H, W),M and a Signature
Step3: σ = (T1, T2, T3, c, sα, sβ, sx, sδ1, sδ2)
Step3: Output: True or False.
Step4: Begin
Compute the following values
R1 = sα. U – c. T1
R2 = sβ. V – c. T2
R3 = (e(T3,W)/e(P,P))c e(T3,P)s
x e(H,W)
R4= sx .T1 - sδ1 .U
R2 = sx .T2 - sδ2 .V
Step5: if c = f(M, T1, T2, T3,R1, R2, R3, R4, R5)
Step6: Return True
Step7: Else
Step8: Return False
Step8: End.

Algorithm 3: Revocation Verification

This algorithm is used to verify active users in the group. Cloud service provider verifies the revocation list before giving access permission to the data.

Step1: Input: System parameter (p, q, r), a group signature M and a set of revocation keys A1.. Ar.
step2: Output: Valid or Invalid.
Step3: begin
Step4: set temp = e = (T1,Q)
e2 = (t2.R)
For i=1 to n
If e (t3-Ai,p)
Return null
Step5: else return temp
Step6: stop

5. EXPERIMENTAL ANALYSIS

In the proposed system, the group manager needs to store the user list and share data. Group manager takes charge of system initialization,

A system with 200 users with an assumption that each user shares 50 files on an average is considered. Then, the total storage of the group manager could be not more than 28.5Kbytes, which is acceptable. Group members need to store only their individual private key which is about 60 bytes. The extra storage overhead to store the file in the cloud is about 248 bytes only.

Therefore, the analysis on the proposed approach shows that the utilization of storage space among different models is low. Thus, it is acceptable for practical usage.

In the proposed scheme, the revocation of user from the group does not increase the computation cost irrespective of the number of revoked users. Revoked users are periodically updated and hence there is no chance of accessing the cloud once they are revoked from the group.

6. CONCLUSION

Sharing secure data in a cloud to multiple users among dynamic groups allows users to share their data with other users in a group without revealing data and identity privacy to the cloud. Additionally, it supports efficient user revocation and new member joining. More specifically, efficient user revocation can be achieved through a publicly available revocation list without updating the private keys of the remaining users and new users can directly decrypt the files from the cloud before their participation by contacting the group manager. Moreover, storage overhead and encryption computation costs are independent of the number of revoked users.

REFERENCES

[1]. P. Junod and A. Karlov, "An efficient public-key attribute-based broadcast encryption scheme allowing arbitrary access policies in Tenth annual ACM workshop on digital rights management. ACM, 2010, pp. 13–24.
[2]. Lam, S.S-zebeni, and L.Buttyan, "Invitation-oriented: Key management for Dynamic groups in an asynchronous communication model," Submitted to 4th International Workshop on Security in Cloud Computing, 2012.
[3]. N. Cao, S. Yu, Z. Yang, W. Lou, and Y. T. Hou, "LT Codes-based Secure and Reliable Cloud Storage Service," in the Proceedings of IEEE INFOCOM 2012, 2012, pp. 693–701.
[4]. B. Wang, B. Li, and H. Li, "Knox: Privacy-Preserving Auditing for Shared Data with large Groups in the Cloud," in the Proceedings of ACNS 2012, June 2012,
[5]. B. Libert and D. Vergnaud, "Unidirectional chosen-ciphertext secure proxy re-encryption," Information Theory, IEEE Transactions on, vol. 57, no. 3, pp. 1786–1802, march 2011.
[6]. S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving Secure, Scalable, and Fine-Grained Data Access Control in Cloud Computing," Proc. IEEE INFOCOM, pp. 534-542, 2010.
[7]. Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S. S. Yau, "Dynamic Audit Services for Integrity Verification of Outsourced Storage in Clouds," in the Proceedings of ACM SAC 2011, 2011,
[8]. H. Abu-Libdeh, L. Prince-house and H. Weather-spoon, RACS: a case for cloud storage diversity, ACM, 2010, pp. 229-240.
[9]. Taka-bi , H.; Joshi, J.B.D.; Ahn, G.; , "Security and Privacy Challenges in Cloud Computing," Security & Privacy, IEEE, vol.8, no.6, pp.24-31, Nov-Dec.2010. doi:10.1109/MSP.2010.186.
[10]. Kamara, Seny and Lauter, Kristin, Cryptographic cloud storage, FC'10 Proceedings of the 14th international conference on Financial Cryptography and data security, pp.136-149, 2010.

To Improvise and Design a Cooperative Secure Data Possession Scheme for Integrity Verification for Data Stored on Cloud

Nisha Singh, Dr. Suneetha K.R.
M.tech, Dept of CSE, BIT Bangalore

***Abstract*-- Cloud is emergent field both in terms of high level computing as well as storage. Cloud provides higher computing facilities by clubbing large number of resources online. The cloud storage is concept of storing data online such that data can be accessed by the user at any location and at any hour. Cloud storage has established its importance as organizational as well as individual facility. The storages available with an organization or an individual are usually in form of physical drives which need to be carried to place where they are meant to be used or shared. Cloud storage provides with an opportunity to access the data from anywhere by storing data remotely and making it available for online access. Due to additional overhead involved organizations choose to outsource their data to a cloud service provider. This gives rise to issue of data security in clouds; this includes issues relating to confidentiality, integrity and availability of the data. For tackling this various schemes have been proposed under provable data possession which challenges the server for possession of data that it claims to have. In this paper we propose a cooperative secure data possession scheme which provides an additional layer of security while providing all features of previous data possession scheme.**

Index Terms -- Multi-cloud storage, Web Server based Multi-cloud, Advanced Cooperative Provable Data Possession, Security in clouds.

I. INTRODUCTION

This paper is conceptualized on the need of security of owner's data which is outsourced at the multiple cloud based web servers. For outsourcing the data the most preferable technology is the cloud. Cloud has services on the basis of pay per use. As data owner outsources data to the CSP, the issue of the security arises because misuse of the outsourced data can be done by CSP itself. There are many ways to prevent misuse of data at the small scale like cryptography.

Security is nothing but the intersection of confidentiality, integrity and availability, failing of anyone parameter to achieve leads to data to become vulnerable. Here all the parameters are assessed but the focus is mostly on the issue of integrity. Confidentiality can be achieved through the authentication while the availability can be achieved through the redundant storage of the data on multiple cloud based web servers.

We first consider integrity, which can be achieved by the technique called PDP. In this technique, the data owner challenges to the CSP for providing the guarantee of the integrity of the outsourced data. The challenge is in the form of query. Query contains some credentials of the uploaded data calculated before uploading. Then the CSP calculates the credentials according to the challenge of the data owner in the form of response, the response is given to the owner. Owner crosschecks the original credentials with the response, if equality holds the owner is satisfied with the provided service and integrity the outsourced data. As mentioned above is already existing technique as per paper [7]. But the overhead and the role of the Data Owner is reduced by inducing the third actor in the picture as TTP. All the above stuff taken place in communication in between CSP and Data Owner is happens through TTP. Next we add another layer of security on the owner side to provide extra control over the data that the user stores on the multi-cloud storage.

II. RELATED WORK

2.1 Survey

All material An Ateniese, et al [2] proposed a PDP model which supports problems of static files. This model works optimally for static case with constant complexity by the principle of blocks and tags. It gives the private authentication and the linear storage for Data Owner. But, this model has some limitations which are detected by later results as, expensive server computation, no security guarantee for data possession, vulnerable to replay attacks.

A. Juels, et al [3] proposed a POR model which also works on static storage with the constant complexity. This varies with the previous model in the processing on the data. Data is modified by inserting some sentinel blocks in between which are used to verify the integrity of the file by checking the correctness of sentinel blocks. This model has both communication and computation complexities constant. It has some limitations on number of times one can challenge for integrity as sentinels are one time labels. Also, the model depends on the large preprocessing of data.

Shacham and Waters [4] proposed Compact POR model which is an improved version of POR. This model uses homomorphic property to aggregate a proof of challenge with authenticity complexity of O(1) and with

computation complexity of O(t). In this model also some limitations with respect to current models as its solution is for static storage and vulnerable to the leakage of data in verification.

All the above techniques are useful for the static storage only. Ateniese, et al [5] proposed Scalable PDP which is the first technique to give the dynamicity in the outsourcing of data. This model works on the principle of the random oracle model. It takes the pre-computed answers as metadata so it limits on the number of times the updates and the challenges can possible. This model supports only append operation on data and does not provide the in between modifications in the already uploaded data, which also puts limitations on the dynamicity of the model.

C. Erway, et al [6] proposed a DPDP model on the basis of PDP model for dynamic storage of files and also which can be updated online. Even after this modification one can verify the integrity of the file. It uses the skip-lists for maintaining tags and is stored at Data Owner side to avoid replay attacks. This model has computational and communication complexity both up to log(n). The server requires the whole path of data block to access it, because this model does not maintain the numbering to the data blocks. In this case, if the file is too large then both the above complexities are considerable.

Feifei Liu, et al [7] proposed Improved DPDP model which is the improvement over the previous model [6]. It does partition of original file into blocks. Tag is generated for each block and hash value is computed for each tag. These tags are used to verify the integrity of the file by using the skip-lists, and the integrity of the tags is verified by the hash values. In this model computational and communication complexities are reduced from log(n) to constant.

2.2 Motivation

To provide a low cost, scalable, location independent platform for managing client's data. Cloud service providers adopt several distributed file systems for cloud storage systems. These systems provide for storage of data over the cloud. These file systems share some similar features: a single metadata server provides centralized management by a global namespace; files are split into blocks or chunks and stored on block servers; and the systems are comprised of interconnected clusters of block servers. Those features enable cloud service providers to store and process large amounts of data. Hence, it is crucial to offer an efficient verification on the integrity and availability of stored data for detecting faults and automatic recovery. However, they do not provide any mechanisms to check for integrity verification of data as well as the availability of the data.

III. PROPOSED METHODOLOGY

3.1 ARCHITECTURE

Although existing PDP schemes offer a publicly accessible remote interface for checking and managing the tremendous amount of data, the majority of existing PDP schemes are incapable to satisfy the inherent requirements from multiple clouds in following terms:

The existing protocols suffer from the following weaknesses.

- Most of the data possession schemes do not take dynamic nature of cloud into account.
- Most of the existing techniques provide integrity verification only in post-attack scenario.
- Most of the existing techniques have large upload and download times.
- Most of the existing techniques add overhead on user for pre-processing of file to check for integrity.

To address these problems, we consider a multi-cloud storage service as illustrated in Figure 1. In this architecture, a data storage service involves three different entities: Clients who have a large amount of data to be stored in multiple clouds and have the permissions to access and manipulate stored data; Cloud Service Providers (CSPs) who work together to provide data storage services and have enough storages and computation resources; and Trusted Third Party (TTP) who is trusted to store verification parameters and offer public query services for these parameters.

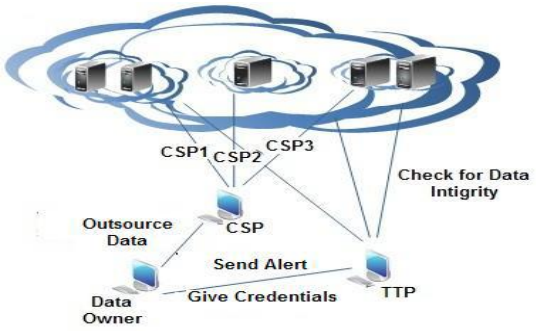


Fig. 1.Verification architecture for data integrity.

In this architecture, we consider the existence ofmultiple CSPs to cooperatively store and maintain theclients' data. Moreover, a cooperative PDP is used toverify the integrity and availability of their stored data in all CSPs. The verification procedure is described as follows: Firstly, a client (data owner) uses the secret key to pre-process a file which consists of a collection of .. blocks, generates a set of public verification information that is stored in TTP, transmits the file and some verification tags to CSPs, and may delete its local copy; Then, by using a verification protocol, the clients can issue a challenge for one CSP to check the integrity and availability of outsourced data with respect to public information stored in TTP. We neither assume that CSP

is trust to guarantee the security of the stored data, nor assume that data owner has the ability to collect the evidence of the CSP's fault after errors have been found. To achieve this goal, a TTP server is constructed as a core trust base on the cloud for the sake of security. We assume the TTP is reliable and independent through the following functions [8]: to setup and maintain the CPDP cryptosystem; to generate and store data owner's public key; and to store the public parameters used to execute the verification protocol in the CPDP scheme. Note that the TTP is not directly involved in the CPDP scheme in order to reduce the complexity of cryptosystem.

3.2 CSDP Algorithm

Step 1: Input file F

Step 2: Using Hash Index Hierarchy concept to split the file.

Step 3: Split file into number of blocks {m1, m2, m3,….}

Step 4: Generate key [7] for each block
Use user's key(sk) applied for keygen(sk,pk)
Where,
sk= encryption of key choosen by user
pk= generated from sk , $0<pk<10^6$

Step 5: Generate Tag [7] for each block
TagBlock(sk, pk, mi, vi, i) →{Ti, hi}
Compute, $T^*_i = g^{mi} \bmod N$
$H^*_i = H_{k1}(Ti \parallel f(vi) \parallel i)$
Where,
H=cryptographic hash function
f = pseudo random function

Step 6: Store all the credentials in the form of tags and hash on the TTP before uploading the data on clouds.

Step 7: Response calculation by TTP from clouds:
Challenge: Query(c) →chal [7]
i. Input: number of blocks to be challenged,b
ii. k2, k3 → random numbers.
index: $i_j = \pi_{k2}(j)$ for $1 \le j \le b$
Coeff.:$a_j = \pi_{k3}(j)$; π = pseudo random permutation
iii. Output: chal = $\{(i_1, i_2, \ldots, i_b), (a_1, a_2, \ldots, a_b)\}$

Proof: Prove(chal, F) → P [7]
i. Input: Query chal, File F
ii. Search h_{ij}, T_{ij}, m_{ij}For $i_j \in \{i_1, i_2, \ldots, i_b\}$

Use the HVR concept [1] to integrate the responses from multiple clouds.
Homomorphism mapped as, f: P → Q
Where, P and Q are two different groups
Such that,

$f(g1 \oplus g2) = f(g1) \otimes f(g2)$; for all g1, g2 ϵ P
Where,
$\oplus$ = operation in P
$\otimes$ = operation in Q

The above homomorphism concept is used for calculating Homomorphic Verifiable Tags such that, T_i and T_j for messages m_i and m_j respectively, then by combining T_i and T_j get T' for the $m_i + m_j$.

By working on the same principle the HVR can be calculated from different responses from different clouds in CSDP scheme and stored at TTP.
Compute, $M = a_1.m_{i1} + a_2.m_{i2} + \ldots + a_c.m_{ib}$
$h = hi_1 . hi_2 \ldots hi_b$
iii. Output: $P = \{M, h, (Ti_1, Ti_2 \ldots Ti_b)\}$

Step 7: Verify the collective response with the RESPONSE stored at TTP.

Verification:
Verify(s_k, chal, P,Ω) → {Accept, Reject}

i. Compute, $T^* = g^m \bmod N$
$T = T_{i1}^{a1} * T_{i2}^{a2} * \ldots * T_{ib}^{ab}$
Search, $\Omega = \{v_{i1}, v_{i2}, \ldots, v_{ib}\}$ for $i_1 \ldots i_b$
$H = H_{k1}(T_{i1} \| f(v_{i1}) \| i_1) * \ldots * H_{k1}(T_{ib} \| f(v_{ib}) \| i_b)$
ii. Check,
T* ?= T, & H*? = h
Where, T* and H* are nothing but the original credentials.
If both holds, output Accept;
Else, output Reject.
Step 8: If there is mismatch, generate alerts.
Step 9: Otherwise, SUCCESS.

IV. CONCLUSION

In this paper, we presented the construction of an efficient cooperative secure data possession scheme for distributed cloud storage. Based on homomorphic verifiable response and hash index hierarchy, we have improvised a scheme to support dynamic scalability on multiple storage servers. We also showed that our scheme provided all security properties required by zeroknowledge interactive proof system, so that it can resist various attacks even if it is deployed as a public audit service in clouds. In this paper we were also able to give another layer of security including the TTP's so the hacker or attacker needs to accomplish a triple layer breach before reaching the client's data.

In future work the scheme could be improved to use higher mechanisms for security enhancements. Also the scheme is dependent on Hash index hierarchy and it needs to be able accommodate requirements of two-layer architecture.

V. REFERENCES

1] Yan Zhu, Hongxin Hu, Gail-JoonAhn, Mengyang Yu, "Cooperative Provable Data Possession for Integrity Verification in Multicloud Storage", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 23, NO. 12, DECEMBER 2012.

[2] G. Ateniese, R.C. Burns, R. Curtmola, J. Herring, L. Kissner, Z.N.J. Peterson, and D.X. Song, "Provable Data Possession at Untrusted Stores," Proc. 14th ACM Conf. Computer and Comm. Security (CCS'07), pp. 598-609, 2007.

[3] A. Juels and B.S.K. Jr., "Pors: Proofs of Retrievability for Large Files," Proc. 14th ACM Conf. Computer and Comm. Security (CCS '07), pp. 584-597, 2007.

[4] H. Shacham and B. Waters, "Compact Proofs of Retrievability," Proc. 14th Int'l Conf. Theory and Application of Cryptology and Information Security: Advances in Cryptology (ASIACRYPT '08), pp. 90-107, 2008.

[5] G. Ateniese, R.D. Pietro, L.V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," Proc. Fourth Int'l Conf. Security and Privacy in Comm. Netowrks (SecureComm '08), pp. 1-10, 2008.

[6] C. Erway, A. Kupcu, C. Papamanthou, and R. Tamassia, "Dynamic provable data possession", In CCS '09, pp. 213-222, April 24,2012.

[7] Feifei Liu, DavuGu, HainingLu,"An Improved Dynamic Provable Data Possession", Proceedings of IEEE CCIS2011, pp 290-295, 2011.

[8] B. Sotomayor, R.S. Montero, I.M. Llorente, and I.T. Foster, "Virtual Infrastructure Management in Private and Hybrid Clouds," IEEE Internet Computing, vol. 13, no. 5, pp. 14-22, Sept. 2009.