CVR JOURNAL OF SCIENCE & TECHNOLOGY

Vol. No: 3, DECEMBER 2012 ISSN 2277- 3916

CVR COLLEGE OF ENGINEERING

In Pursuit of Excellence

PATRONS

CVR JOURNAL
OF
SCIENCE & TECHNOLOGY

CVR COLLEGE OF ENGINEERING

(An Autonomous College affiliated to JNTU Hyderabad)
Mangalpalli(V), Ibrahimpatan(M),
R.R.District, A.P. – 501510
http://cvr.ac.in

EDITORIAL

The Third Volume of CVR Journal of Science & Technology, a Bi-annual Journal with ISSN number published by our college, is being brought out with 18 papers from various areas of specialization in Engineering and Science. The papers have been selected from several papers received, after careful review by the members of the Editorial Committee. The breakup of papers among the branches is as follows:

ECE - 5, CSE - 4, EEE - 2, EIE -2, H&S- 2, IT -2 and Management -1.

Research in Science and Engineering, complementing each other, is resulting in the current fast paced developments taking place on a global scale, affecting the day to day life the world over. Engineers have a crucial role in this context and the quality of Engineering Education becomes particularly important in preparing the students for challenges to be faced by them in their future career. Teachers have a major role to play for this purpose. Research culture in the college contributes in no small measure to equip the teachers with the appropriate attitudes and skills for pursuing research in the areas of current interest. Further, topical research by teachers complements their teaching activity, thereby improving the quality of teaching and project guidance, and thus ensuring that the outgoing students will be competent engineers.

The close link between Science and Engineering is well recognized. A sound background in Science helps engineers in making judicious decisions in the design and implementation process of solving real world problems in an optimum way. Likewise proper utilization of the improved and sophisticated tools, provided by progress in engineering, should help in optimizing the pursuit of scientific research. CVR Journal of Science & Technology continues with the tradition of making the best use of the complementary nature of Science and Engineering, and strengthening the research base of the college.

It will be our earnest effort to continue to maintain and improve the quality of the Journal with active support of the Management of the College.

K.V. CHALAPATI RAO

CONTENTS

CVR Journal of Science & Technology, Volume 3, December 2012 ISSN 2277-3916

Generation Of 3D City Model From Multi-Sensor Geo-Spatial Image Data And It's Visualization

V. NaveenChandra[1] and Dr. N. V. Rao[2]
[1] CVR College of Engineering, Department of CSE, Ibrahimpatan, R.R.District, A.P., India
Email: vnc526@gmail.com
[2] CVR College of Engineering, Department of CSEIT, Ibrahimpatan, R.R.District, A.P., India
Email: nvr@ieee.org

***Abstract*— Generation of 3D city models has been evolving to be one of the most relevant and active research topics as many application areas require them. Planning of Telecommunications, Urban development, Managing disasters, Prevention of Hazards, City design, developing TV games, etc., are some of the application areas where 3D city models are precisely and accurately needed. Several concepts from Geo-Information System (GIS) were taken for transmitting traditional 2D models to 3D City Models.**

The paper describes the Generation of 3D city model in a Semi-Automated Approach. Different multi-sensor data/Images were studied and analyzed, and the images were pre-processed using image processing techniques. The 3D city model generation in semi-Automated approach involves development of prototype software which eventually can be utilized in the Automatic generation of 3D city model. The entire prototype software is designed and developed using MATLAB platform. The generated 3D city model is represented or visualized using suitable visualization techniques.

***Index Terms*— GIS, Photogrammetry, LIDAR, GPU, DEM, DSM, DTM, Aerial photography.**

I. Introduction

3D city model graphically simplifies complicated concepts which are difficult to visualize. The technical standards and commercial possibilities are high in 3D representation, as in today's world people are smart enough to use 3D city model in tremendous technical, environmental and commercial aspects such as marketing, tourism, city planning, traffic management, navigation, noise propagation and much more.

3D city model consists of buildings, streets, terrain, vegetations and other man-made objects. Modeling of buildings is the main workload and is clear that photogrammetry is able to provide a means to collect the required 3D information. Generation of 3D city model is broadly classified into two large classes, namely Automatic systems [1] and Semi-Automatic systems [2]. In Automatic systems, 3D city models are generated without any human interaction. The system itself detects all the objects such as roads, trees, grass, rivers, hills, etc. whereas in Semi-Automatic approach user manually selects the buildings, towers, trees, etc and based on the selection 3D city model is generated.

In this paper we have presented 3D city model developed in semi-automated approach. The paper focus on research work for generation of 3D city model from geo-spatial data collected from multiple sources. Initially, different multi-sensor data and images were analyzed and pre-processed. Then, 3D city model was constructed and is represented using suitable visualization techniques. The prototype is designed and developed using MATLAB platform and is optimized using GPU (Graphical Processing Units) based computing techniques. The prototype software developed here will be eventually utilized to generate automatic systems.

Visualization through visual imagery has been an effective way to communicate both abstract and concrete ideas since the dawn of man. The tool used here for visualization of 3D city model is ParaView, an open-source, multi-platform data analysis and visualization application. 3D city buildings generated in this approach is stored as a grid file in 'VTK' (Visualization Tool Kit) format. VTK is an open-source, freely available software for 3D computer graphics, image processing and visualization.

Geo-Spatial Image data refers to different kinds of data such as Aerial Photography, Topographic Images, Digital Surface Model (DSM) files, Digital Terrain Model (DTM) files, Digital Elevation Model (DEM) files, LIDAR (Light Detection and Ranging) data, Satellite Imagery and TLS (Three Line Scanner) Image information.

A. Abbreviations and Acronyms

GIS (Geographical Information System):GIS is a system of systems collecting, processing, and analyzing Spatio-Temporal information regarding earth features [6]. It involves people preparing data, the system (I/O devices, computing platforms and networks) and Users using the system.

DEM (Digital Elevation Model):DEM is a digital model for 3D representation of terrain's surface.

DSM (Digital Surface Model):DSM represents the earth's surface and includes all objects in it.

DTM (Digital Terrain Model):DTM represents the bare ground surface without any objects like plants and buildings.

LIDAR (Light Detection and Ranging):It is an optical remote sensing technology that can measure the distance to, or other properties of a target by illuminating the target with light [13], often using pulses from a laser.

TLS (Three Line Scanner):Three Line Scanner (TLS) is an optical sensor for aerial survey. TLS is composed of three linear CCD arranged in parallel, and it can acquire three images of each direction (Forward, Nadir and Backward) at the same time. Orienting them on an aircraft perpendicularly to flight direction, and scanning a ground plane, a treble stereo image of a ground object can be acquired [3].

GPU (Graphics Processing Unit):A graphics processing unit or GPU (also occasionally called visual processing unit (VPU)) is a specialized electronic circuit designed to rapidly manipulate and alter memory in such a way so as to accelerate the building of images in a frame buffer intended for output to a display. GPUs are used in embedded systems, mobile phones, personal computers, workstations, and game consoles.

CSG (Constructive Solid Geometry):Geometric methods to build roofs of buildings such as Flat, Desk, Gable and Hip roofs.

Apart from the above terms and definitions, some of the keywords include Aerial photography, Satellite Imagery, Topographic Images, Terrestrial images, Stereo Image Data, etc., comes under Geo-Spatial data.

II. CONSTRUCTING 3D CITY MODEL

3D city models are digital representations of the Earth’s surface and related objects belonging to urban areas (like cities, factories, buildings, etc). MATLAB supports various forms of Image formats. The Input Image can be a TIF, BMP, JPEG, PNG, GIF, etc,. MATLAB processes it and can produce preferable output. A GUI is designed for processing of images and to generate the expected model.

A. Selection of Buildings

Consider any aerial image and select the corner points of buildings. Based on the markings, draw the outlines of buildings to verify whether the lines are interacting with the geometrical shapes of buildings. If not then erase the old markings and new ones are replaced.

B. Mapping the boundaries with building edges

As we get the exact outlines of buildings, we will extract the boundary lines of buildings and are mapped with the buildings edges that we get by implementing edge detection techniques such as sobel’s method, canny’s method and robert’s method. These algorithms give us fine line extraction [2] of buildings which is the input for generating 3D city model.

Figure 1. Markings of buildings and their outlines

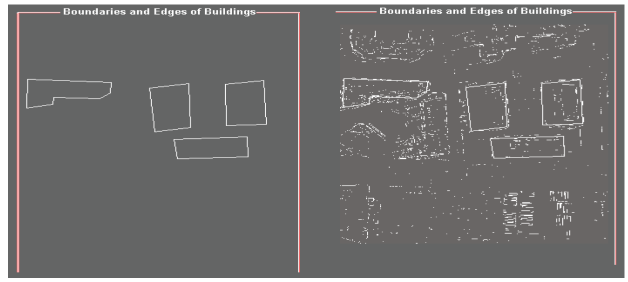


Figure 2. Mapping boundaries and edges (Robert's edge detection)

C. Generating 3D city model

Digital Elevation Model provides the height data for a particular area on which we are generating 3D model. For the time being assuming a particular height for all buildings, 3D city model is generated using fine line boundaries which are the efficient input. 2D image is kept as a surface and on top of that 3D projections were drawn according to the boundaries obtained. 3D model from a traditional 2D image is successfully completed.

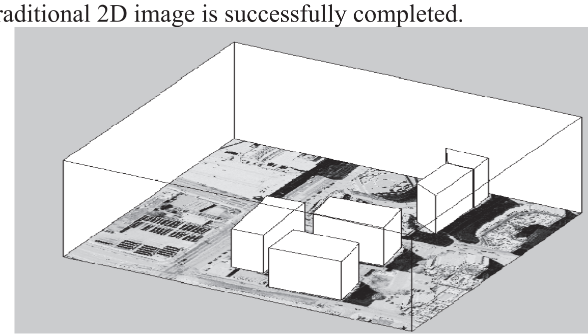


Figure 3. Generated 3D city model

III. GENERATING A GRID FILE (VTK) AND VISUALIZATION OF 3D CITY MODEL

A. Grid file generation

Assume a 3D grid with our 2D image as its base with some respective value of height. As we have developed 3D city model by developing the building blocks on the top of 2D image, so the points at which the buildings are touching the 3D grid were considered and stored in an external file (‘.vtk’ file). VTK has an extensive information visualization framework, having a suite of 3D interaction widgets, It supports parallel processing, and integrates with various databases on GUI toolkits such as Qt and Tk. VTK is a cross-platform and runs on Linux, Windows, Mac and Unix platforms.

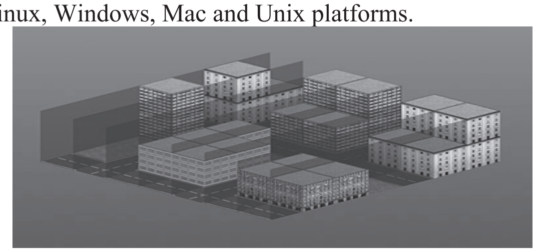


Figure 4. City model in 3D Grid

B. Visualization using ParaView

The generated ‘vtk’ file can be visualized using visualization tool ParaView. ParaView is an open-source, multi-platform data analysis and visualization application. ParaView users can quickly build visualizations to analyze their data using qualitative and quantitative techniques. The data exploration can be done interactively in 3D or programmatically using ParaView's batch processing capabilities. When the 'vtk' file is opened in ParaView the buildings of City model are visualized. This 'vtk' file is subsequently used for various applications.

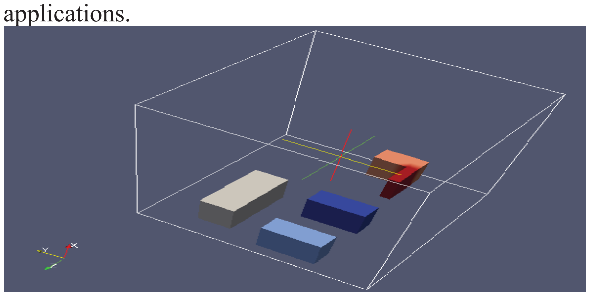


Figure 5. Visualizing 'vtk' grid file using ParaView

CONCLUSIONS

Generation of 3D city model from Aerial images in Semi-Automated approach is successfully done. The prototype software which is built here will be utilized for Automatic Generation of 3D city Model. The 'vtk' file generated is the efficient input for effective visualization of 3D city Model. The external file (vtk) is portable and hence 'vtk' file generated for any area can be easily carried and it can also be used in many applications such as traffic management, tourism, urban planning, marketing, telecommunication planning, disaster management, etc.

In the fully automated approach, the system itself detects the objects such as buildings, trees, roads, etc. in an image. In semi-automated approach the user manually selects the objects in an image. So, except the detection part remaining procedure for generating 3D city model in automated and semi-automated systems is the same. Hence, the modules developed in our approach can be used in developing automated 3D city model generation.

A. Scope for further work

- To have a real world scene like looking of 3D city model, Pattern Matching is to be done.
- Buildings are classified into two kinds. Convex shaped and Concave shaped structures. For Concave shaped buildings like 'U' shaped, 'V' shaped, etc., Semi- automated approach needs more number of interactions from user. So implementing Delaunay triangulation method may reduce interactions per building.
- Delaunay triangulation draws extra lines i.e. lines that do not belong to a building. We should remove those lines. Automatically removing them will be an efficient work.
- Coming to visualization we have developed grid values of buildings that are rectangular shaped only. Generating grid for all regular and irregular shapes (triangular, spherical, trapezium, etc.) looks very effective and therefore very sophisticated city models can be easily built from this 'vtk' grid file.
- 3D models of buildings are developed assuming same height for all buildings. If height data such as DEM files are available we can generate more effective and realistic 3D City Models.

ACKNOWLEDGMENTS

The present work is based on the M. Tech project work done by the first author at ANURAG (DRDO), Hyderabad . The authors are grateful to the Director, ANURAG and Mr. Dilip Kumar Dalei, Mr. Ravi kumar and Mr. B V H K Nanada, Scientists in ANURAG, DRDO for their permission to use available resources and their guidance throughout this work. The first author is thankful to Prof. L. Siva Reddy (Vice principal & HOD, CSE Department) and to the Principal for providing all the facilities during his studies at CVR college of Engineering.

REFERENCES

[1] Y. Takase, N. Sho, A. Sone, K. Shimiya. Automatic Generation of 3D city Models and related Applications. International Archives of photogrammetry, remote sensing and Spatial Information Sciences, Vol. XXXIV-5/W10.

[2] Zuxun Zhang, Jun Wu, Yong Zhang, Yongjun Zhang, Jianquing Zhang. Multi-View 3D city model generation with Image sequences. International Archives of photogrammetry, remote sensing and Spatial Information Sciences, Vol. XXXIV-5/W12.

[3] "Study on making City Model with Image and Laser Range Data" *(Paper Presented at the 22nd Asian Conference on Remote Sensing, Singapore)* Masafumi, Nakagawa.

[4] Yoshiaki, Kagawa, 2001. Automatic acquisition of 3D city data with air-borne TLS (Three Line Scanner) and Laser scanner.

[5] Brenner, C., Haala, N. and Fritsch, D, 2001. Towards fully automated 3D city model generation. In Automatic Extraction of Man-Made Objects from Aerial and Space Images III. 2001.

[6] Grun, A., Wang, X., 1998. CC Modeler: A Topology Generator for 3-D City Models. ISPRS Commission IV Symposium on "GIS - Between vision and application", Stuttgart/Germany, IAPRS. Vol. 32, Part 4, pp. 188-196.

[7] Construct 3D City Model by Multi-Sensor Data. Fei Deng, Zuxun Zhang, Jianqing Zhang. Remote Sensing School, Wuhan University, Wuhan, China-dengf-wh@163.com

[8] Henricsson O., Bignone F., Willuhn W., Ade F., Kübler O., Baltsavias E., Mason S., Gruen A., 1996. Project AMOBE: Strategies, Current Status and Future Work. Paper presented at the 18. ISPRS Congress, 9. - 14. July, Vienna, Austria. In IAPRS, Vol. 31, Part B3, pp. 321 - 330.

[9] Maas, H.-G., 1999. Closed solutions for the determination of parametric building models from invariant moments of airborne laser scanner data. ISPRS Conference 'Automatic Extraction of GIS Objects from Digital Imagery, München/Germany.

Genetic Algorithm Based Dynamic Jobs Scheduling in grid computing

Pritibahen Sumanbhai Patel[1] and B.Jyoshna[2]
[1] CVR College of Engineering, Department of CSE, Ibrahimpatan, R.R. District, A.P., India
Email: priti.patel0209@gmail.com
[2] Nishitha College of Engineering and Technology, Department of CSE, Hyderabad, India
Email: bmjyoshna@gmail.com

***Abstract*-A computational grid is a large scale, heterogeneous collection of autonomous systems, geographically distributed and interconnected by low latency and high bandwidth networks. The sharing of computational resources is a major aspect of grids. Scheduling is a key problem in emergent computational systems, such as Grid and P2P, in order to benefit from the large computing capacity of such systems. Our approach is to dynamically generate an optimal schedule so as to complete the different tasks in a minimum period of time as well as utilizing the resources in an efficient way. There are so many approaches for scheduling like Genetic Algorithm (GA), Simulated Annealing (SA), Ant Colony optimization (ACO) and Particle Swarm Optimization (PSO) Algorithm. In this paper, We would like to present Genetic Algorithms (GAs) based schedulers for efficiently allocating jobs to resources in a Grid system. We would also like to implement GAs for designing efficient Grid schedulers when makespan is minimized. Our GA-based schedulers are very fast and hence they can be used schedule jobs arrived in the Grid system.**

***Index Term* - Genetic Algorithm, Makespan, Minimum completion time, Fitness.**

I. INTRODUCTION

Grid computing has emerged as an important field, distinguished from conventional distributed computing by its focus on large-scale resource sharing, innovative applications and high-performance orientation. In grid computing scheduling is challenging job. So we used GAs for designing efficient Grid schedulers when makespan is minimized. The GA operation is based on the Darwinian principle of "survival of the fittest". It implies that the fitter individuals are more likely to survive and have a greater chance of passing their good genetic features to the next generation. In genetic algorithm, each individual that is a member of the population represents a potential solution to the problem. GA starts with initial population of individuals (chromosomes). Each individual is evaluated using fitness function to produce a value known as goodness of the solution. Then a new population is generated by selecting best individuals from the current population and applying crossover operator to produce new offspring which would inherit good features of parents. Then each offspring is mutated in order to prevent GA to be trapped in local optima. Best individuals among current population and new population are carried forwarded in the next generation. The process is repeated until stopping condition met and best solution in the current generation is returned. We have used Genetic Algorithm based approach forour paper because GA can search for optimal/nearly optimal solution for scheduling quickly. It is well understood and applicable to many real life problems. GA can easily be combined with other meta-heuristic approaches for multiple objectives.

II. LITERATURE SURVEY

The existing approach for grid scheduling implemented with conventional algorithm techniques may give optimal solution but not in reasonable amount of time & the literature shows several limitations. These are : Algorithms are studied using simulation, mostly static algorithms which assume that all information is known in advanced, do not react to dynamism involved in the typical grid environment and The performance of these algorithms has been studied for small sized problems only. So in this paper, we proposed to implement a grid scheduler which will address all of the above problems. It is based on Genetic algorithms which gives optimal/nearly optimal solution quickly. It uses dynamic information received from Grid Information System to determine optimal/ nearly optimal solution. It can work with larger sized problems.

III. PROPOSED SYSTEM DESIGN

I used genetic algorithm to find optimal/nearly optimal schedule when makespan is minimum which efficiently utilize the resources. Proposed GA can quickly search solution space in parallel to find optimal/nearly optimal solution in very less time. It uses dynamic information received from Grid Information System to determine optimal/ nearly optimal solution. It can work with larger sized problems. We are going to present a job scheduling algorithm which can perform well.

IV. PROBLEM FORMULATION

Our GA is based on Expected Time to Compute (ETC) Model. An ETC for any job j on any resource (machine) r is expected execution time of job j on r if j is scheduled on r. The problem for grid scheduling consists of following:

- n – the number of jobs to be schedule at particular instance of time. Any job has to be processed entirely in unique resource.

- m – the number of heterogeneous resources(machines) available in the Grid for an execution of a given set of jobs
- N = $\{j_1, \ldots, j_n\}$ a set of n jobs
- R = $\{r_1, \ldots, r_m\}$ set of available m resources.
- The workload Wi of each job i.
- The computing capacity CC_r of each resource (in millions of instruction per second) r.
- The expected time to compute ETC matrix of size n×m(number of jobs * number of resources).ETC[j][r] indicates the expected execution time of job j in resource r.

I considered the scenario in which jobs submitted to the Grid are independent and are not preemptive.

A. Fitness of a Schedule

We used uni-criteria optimization case for computing optimal/nearly optimal schedule of a set of jobs on a set of heterogeneous resources. The fundamental criterion is that of minimizing the makespan.

B. Makespan

The time when latest job finishes. It is calculated as follows:

$$makespan = \min_{S_i \in Schedules} \left\{ \max_{j \in N} F_j \right\} \ldots\ldots\ldots (1)$$

In eq.(1) F_j denotes time when job j finalizes, Schedules denotes the set of all possible schedules and N denotes the set of all jobs to be scheduled. The goal of scheduler is to maximize resource utilization and minimize makespan. Completion time of machine i is denoted by completion[r] and it is expressed as a total time needed for the resource r to finalizing its previously assigned jobs and jobs which are actually scheduled to this resource. We can compute ETC and completion time completion[r] for resource r as follows:

$$ETC[j][r] = \frac{W_j}{CC_r} \ldots\ldots\ldots (2)$$

$$completion[r] = ready_r + \sum_{\{j \in N | Schedules[j]==r\}} ETC[j][r] \ldots\ldots\ldots (3)$$

Where,

ETC[r][j]=expected time to compute job j on resource r.
W_j=workload of job j
CC_r=computing capacity of resource r.
completion[r]=completion time for resource r.
$ready_r$=time when resource r finishes previously assigned jobs to it.

The makespan of eq.(1) can be redefined as the maximal completion time and can be calculated as follows:

$$makespan = \max\{completion[r] | r \in R\} \ldots\ldots\ldots (4)$$

A criteria makespan can be integrated in several ways to establish the desired priority among them. In the multi-objective optimization two fundamental models are the hierarchical and the simultaneous approach. In hierarchical approach, the optimization criteria are sorted by their importance. The process starts by optimizing most important criterion. When further improvements are not possible, the second criterion is optimized while keeping optimized value of first important criterion unchanged. In grid scheduling, makespan may be considered as most important criterion. We used simultaneous approach to compute objective function or fitness function.

Fitness = 1 / makespan … … … … … … … … (5)

V. Overall System Architecture

We implemented GA based grid scheduler that maximizes resource utilization by minimizing makespan. It also determines schedules based on the current resource information (dynamic and static information). And hence can easily react to dynamism involved in grid environment. Overall system architecture shown in fig.1.

We designed our system in 3 major modules.

A. Monitoring & Discovery Service (MDS) Module

This module is used to discover the new grid resources and to monitor already discovered resources. When MDS process starts first time it reads /var/grid resources file to get list of the resources available initially. It also creates a thread to periodically poll already discovered grid resources to get current information about each of these grid resources. The information includes static information about resources such as processor family/architecture, number of CPUs/resource, CPU frequency, total RAM, total swap area etc., and dynamic information such as resource computing status busy/free, resource up/down status, free RAM, load, number of free CPUs etc. It also periodically receives resource information from grid resources. This information is sent to manager process as well as GA based grid scheduler as and when needed. GA scheduler uses current resource information to compute optimal/nearly optimal solution to assign jobs to resources. It also receives update information from manager process and updates its data structures accordingly.

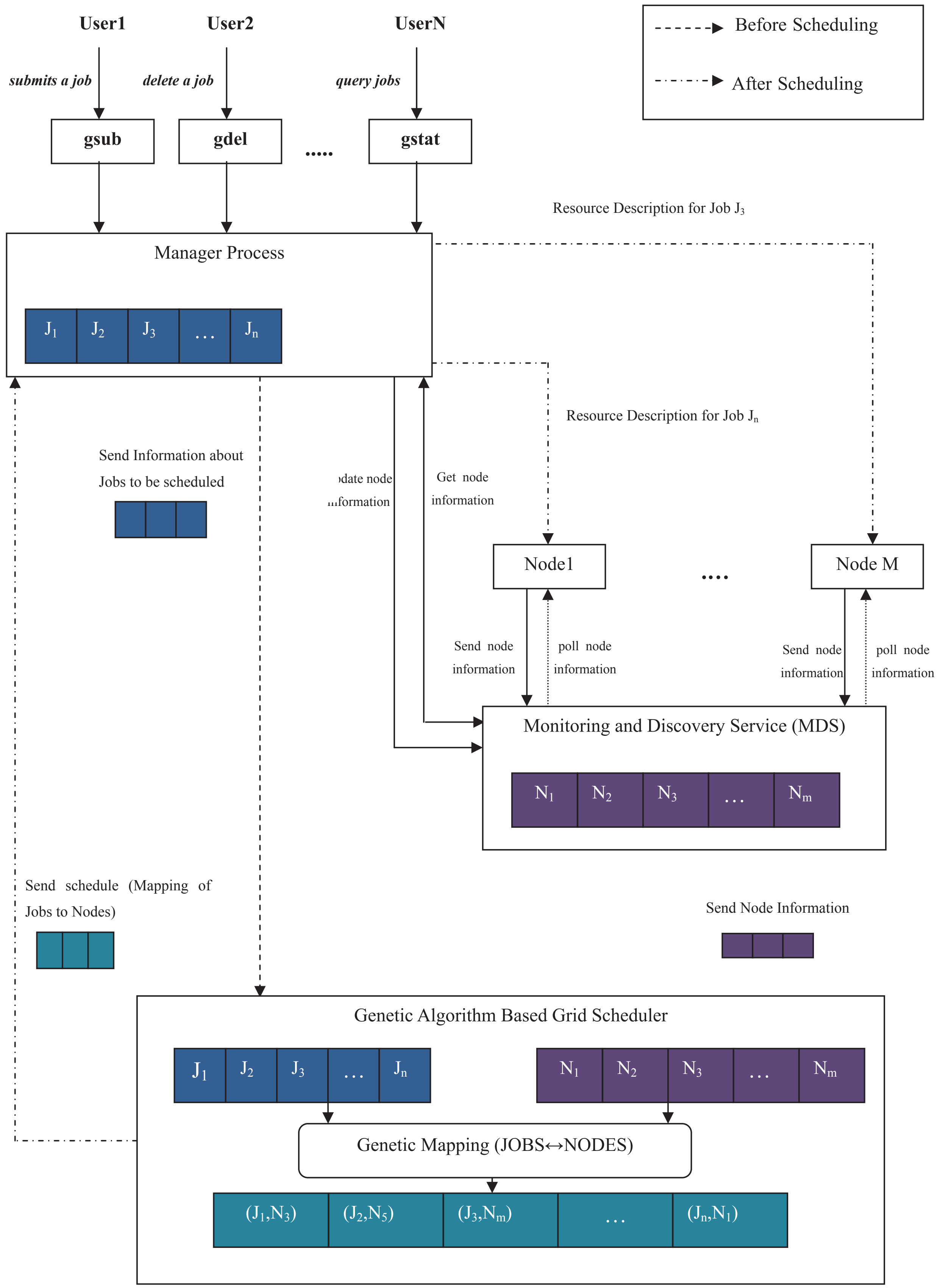


Figure 1. Proposed system design

B. Manager Module

This module is the central part of our implementation. It receives commands from users. It implements following functions.

- Command processing & Scheduler invocation
- Job queue management
- Job management
- Job monitoring

The Manager receives command requests (such as submit a job, query jobs, delete a job) from users as shown in figure 1. When a user submits a job using gsub command, it sends job submission request to manager. When manager receives a job submission request, unique job id is generated for a job and its description is appended to job queue. If a command request is to query jobs(gstat), it simply loop through job queue and send information such as job id, job status, job name, job executable, assigned resource if it is already scheduled etc. If command is to delete a job (gdel) and job is scheduled then job management components forward request to gatekeeper of the assigned host to clean the job. Once the job is deleted on the resource, it will be removed from the job queue otherwise an error is reported. This component periodically checks if there are unscheduled jobs in the job queue. If there are some jobs, it connects to GA Grid scheduler, send information about jobs to GA grid scheduler and wait for optimal/nearly optimal mapping of jobs to suitable resources from the scheduler. Once it receives, a optimal/nearly optimal schedule from scheduler, for each (job, resource) pair in the schedule, it submits to local resource manager for execution purpose.

C. Scheduler Module:

This module uses Genetic Algorithm to find optimal/nearly optimal solution by minimizing makespan. It receives information about list of jobs from manager and information about available resources from MDS server. It then creates initial population of k schedules using Minimum Completion Time heuristics. It then evaluates the current population by computing fitness function for each of k. It then creates a new population by repeating selection, crossover, and mutation and assignment steps until the size of new population becomes k. It then evaluates the new population and carries forward best schedules of the current population as well as the new population in the next generation in order to get optimal/nearly optimal solution quickly. The algorithm evolves generation by generation until termination criteria met. The Scheduler then return best schedules in current population. This schedule will then be sent to manager. Manager submits this job description to the assigned resource.

VI. System Design

This section presents actual design of our system which is Job scheduler using Genetic Algorithm in grid computing. Dynamic task scheduling using Genetic Algorithm in a computational grid, resources are shared by many users, who submit their applications concurrently. We implemented Genetic Algorithm based Grid scheduling using following steps.

A. Schedule encoding

We used direct representation to encode each possible schedule in a chromosome. We used array chromosome of n(number of jobs) integer to represent a chromosome(a schedule) as shown in Figure 2. Chromosome[j] represents the resource number where job j is scheduled.

Job No:	1	2	3	4	5	6	7
Resource No:	4	2	7	6	3	5	1

Figure 2. Encoding of a schedule (a chromosome)

B. Generation of Initial population

In GA, initial population is usually generated randomly. But to guide the searching process and to get optimal/nearly optimal solution in fewer generations, several problem specific heuristics may be used such as Min-Min, Minimum Completion Time (MCT) etc. We used MCT heuristics to guide a searching process for finding optimal/nearly optimal schedule quickly in fewer generations. In the MCT heuristic, each job is assigned to the resource where job completes in minimum time. Jobs are considered for allocation at random.

C. Compute Fitness function

The scheduler aims to maximize resource utilization by minimizing makespan. Good chromosomes have higher fitness values. The fitness of each chromosome (schedule) is computed using equation (5).

Selection operator: Selection operator is used to select parents to which crossover operator is applied to produce new offspring. In general, selection is directly proportional to the fitness of chromosomes. Several selection methods exist to select chromosomes for crossover such as linear ranking, roulette wheel selection etc. We used roulette wheel selection technique to select good schedules to produce new offspring. In roulette wheel selection method, the probability that a chromosome selected is directly proportional to its fitness value. Higher the fitness, higher chances the chromosome will be selected. In this method, each schedule or chromosome gets portion on the roulette wheel according to its fitness value. Chromosomes with higher fitness value get larger slice on roulette wheel. Selection is done by spinning a roulette wheel. Since fittest schedule has larger portion on the roulette wheel, they will have higher chance of being selected. Circumference of roulette wheel represents the total fitness of all chromosomes. Pseudo code for roulette wheel selection method is shown in Figure 3. The roulette wheel selection of among 4 chromosomes is shown in Figure 4. Chromosome 3 has higher chance of getting selected as shown in Figure 4.

```
RouletteWheelSelection()
{
total_fitness=0.0;     running_sum=0.0;
for each chromosome k in a current population
              total_fitness=fitness(k);
r=select random number r in the range [0,total_fitness-1
for each chromosome k in a current population
           running_sum=running_sum+fitness(k);
             if(running_sum >= r)
                 return(k);
}
```

Figure 3. Pseudo code Roulette Wheel Selection

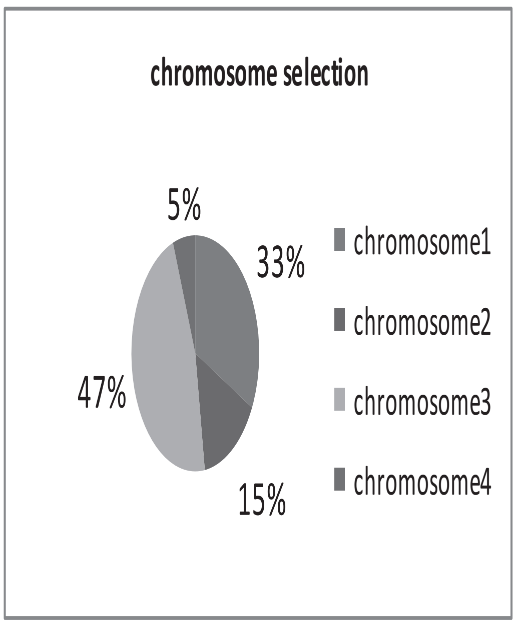


Figure 4. Roulette wheel selection among 4 chromosomes

D. Crossover operator

With crossover operator, two selected parent chromosomes can interchange their genes and produce new offspring (children). The aim is to obtain better quality solution and explore a new region of solution space that has not been yet explored. One may use several different types of crossover such as one-point crossover, two-point crossover, uniform crossover etc. We used one-point crossover operator to produce offspring schedules. In this method, first, random crossover point between 1 and n(number of jobs) is selected, and then first parts of two parents are interchanged to produce two offspring(schedules). Same way, exchanging second parts of two parents to produce two new offspring (schedules) which are same as those produced by exchanging first parts. One point crossover is explained in Figure 5.

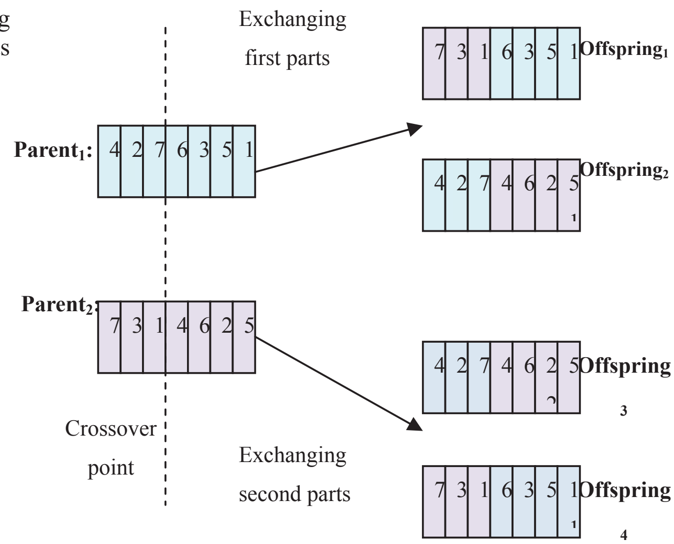


Figure 5. One-point crossover operation to produce 4 offspring schedules

E. Mutation operator

Mutation randomly changes gene(s) to different values. It is used to provide diversification by changing some gene(s) randomly and thereby prevent GA search process getting stuck in to local optima. There are several types of mutation such as move, swap etc., applied to a schedule. We used move mutation which randomly selects a job in a schedule (a chromosome) and assign it to another machine as shown in Figure 6.

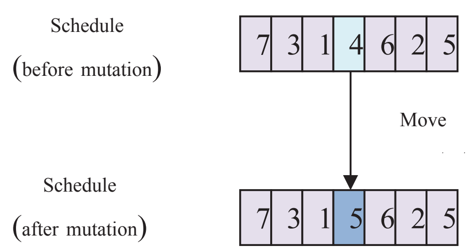


Fig 6: Mutation operation

F. Replacement operator

Replacement operator determines which of the chromosomes (schedules) survives in the next generation. Two kinds of replacement usually used to carry forwards chromosomes to next generation (a) Generational replacement (b) Partial replacement. In a generational replacement, the current population is entirely replaced by new population while in partial replacement worst chromosomes in a current population are replaced by good chromosomes of new population. We used partial replacement strategy in which k best chromosomes from combined current and new population are carried forward to the next generation. First, fitness function is computed for each offspring. Let CP(t) be the current population in generation t and NP(t) be new population in generation t, then current population of next generation t+1 will be

CP (t+1)=k best schedules from (CP(t) U NP(t))

G.. Termination Criteria:

Termination criteria could be:

(i) Maximum number of generations or iterations: the genetic search process is terminated after fixed number of generation.

(ii) Number of iterations without improvement: the optimization process is terminated after some fixed number of iterations without any improvement.

We used (i) termination criterion for our genetic algorithm based grid scheduler in which search process terminates after 300 generations.

If termination criterion is not satisfied goto step 3 and repeat the process.

In general, this genetic search process can be summarized as follows:

GAGridScheduling() {

1. ENCODING: Represent a schedule(a chromosome) using array of n(numof jobs) integer chromosome such that chromosome[i] represents the resource on which job is scheduled
2. INITIALIZATION: Generate a initial population CP(t=0) of k schedules using MCT(Minimum Completion Time) heuristic.
3. FITNESS: Evaluate schedule in CP(t) using eq. (5)
4. TERMINATION CRITERIA: Check if termination criteria satisfied, if 'yes' return the best solution from current population CP(t).
5. NEW POPULATION: Repeat following steps until size of new population NP(t) becomes k.

   (a) Selection: Select two parents schedules p1 & p2 from CP(t) using roulette wheel method.

   (b) Crossover: With crossover probability p_c perform one-point crossover to produce two new offspring schedules o1 & o2.

   (c) Mutation: With very low mutation probability $p_{m,}$ change the assignment of randomly chosen job to new grid resources in each offspring o1 and o2.

   (d) Assignment: Place o1 & o2 in NP(t)

   NP(t)=NP(t) U{ o1,o2}
6. FITNESS: Evaluate schedule in NP(t) using eq.(5).
7. REPLACEMENT:

   (a) Select k best schedules from CP(t) and NP(t) to carry forward in the next generation. CP(t+1)=k best schedules from (CP(t) U NP(t))

   (b) Increment generation count

   t=t+1

   Goto Step 4

}

Table I. List of grid resources with corresponding computing capacity

Resource No.	Computing Capacity (MIPS)	Existing workload (pending processing in ms)
1	3380	72.88
2	931	43.44
3	2969	69.92
4	3120	97.47
5	3728	47.61
6	1815	32.22
7	3170	22.67
8	2084	46.86
9	2014	26.48
10	3318	46.09

VII. RESULTS & ANALYSIS

For the experimental purpose consider following problem instance consisting of 10 grid resources and 20 jobs. List of grid resources with existing workload is shown in the Table I.

Table II. List of jobs with corresponding workload

Job No	Workload
1	126
2	233
3	759
4	858
5	829
6	255
7	789
8	898
9	547
10	110
11	595
12	394
13	582
14	394
15	908
16	310
17	568
18	530
19	125
20	804

To find out optimal/nearly optimal solution for this problem instance, we tuned our genetic algorithm based scheduler with following parameters.

Number of Generations=300
Size of population=256
Crossover probability (P_c)=0.90
Mutation probability (P_m)=0.0001

We got makespan=26.0183 in generation number 189 and then it retains this value until last generation. So if we reduce number of generations to less then 189, we got makespan=26.6066. The graph of generation numbers vs makespan for this problem is shown in fig. 7 where Y-axis represents makespan values and X-axis represents generation number.

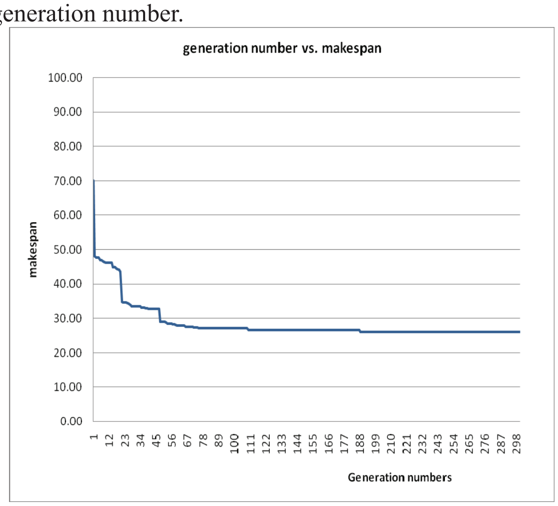


Figure 7. Makespan vs. generation numbers

Conclusion

We presented an extensive study on the usefulness of Genetic Algorithms (GAs) for designing efficient Grid schedulers when makespan parameter is minimized under hierarchic and simultaneous approaches. The experimental study reveals the quality of the proposed GA-based schedulers as compared well to the existing GA-schedulers in the literature. Our GA-based schedulers can be used to design dynamic schedulers. A dynamic scheduler would run our GA in batch mode to schedule jobs arrived in the system since last activation of the scheduler.

As part of our future work we plan to extend focus on workflow based scheduling. Workflow management system allows the user to specify their requirements along with the descriptions of tasks and their dependencies using the workflow specification. Many Grid applications such as bioinformatics and astronomy require workflow processing in which tasks are executed based on their control or data dependencies. It will be integrated with various grid middleware such as UNICORE, LIGEON etc. Study of implementing same scheduler with different heuristics such as min-min, max-min, MET etc. Extending scheduler for multi match making between user's requirement and resource characteristics.

References

[1] Abraham, A. H. Liu, W. Zhang and T. G. Chang, Job scheduling on computational grids using fuzzy particle swarm algorithm, Proc. of the 10th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, B. Gabrys et al. (eds.): Part II, Lecture Notes on Artificial Intelligence 4252, 500507, Springer, 2006.

[2] Abramson, D., R. Buyya and J. Giddy, A computational economy for grid computing and its implementation in the Nimrod-G resource broker, Future Generation Computer Systems Journal, vol.18,no.8, pp.1061-1074, 2002.

[3] Alba, E., F. Almeida, M. Blesa, C. Cotta, M. Daz, I. Dorta, J. Gabarr, C. Le, G. Luque, J. Petit,C. Rodrguez, A. Rojas and F. Xhafa, Efficient parallel LAN/WAN algorithms for optimization,Parallel Computing, vol.32, no.5-6, pp.415-440, 2006.

[4] Buyya, R., Economic-based Distributed Resource Management and Scheduling for Grid Computing, Ph. D. Thesis, Monash University, Melbourne, Australia, 2002.

[5] Buyya, R., D. Abramson and J. Giddy, Nimrod/G: An architecture for a resource management and scheduling system in a global computational grid, Proc. of the 4th International Conference on High Performance Computing, Asia-Pacific Region, China, 2000.

[6] Javier Carretero, Fatos Xhafa, Ajith Abraham. Genetic algorithm based schedulers for grid computing systems. In International Journal of Innovative Computing, Information and Control ICIC International °c 2005 ISSN 1349-4198 Volume 3, Number 6, December 2007.

[7] A. Abraham, R. Buyya, and B. Nath. Nature's heuristics for scheduling jobs on computational grids. In The 8th IEEE International Conference on Advanced Computing and Communications (ADCOM 2000), India, 2000.

[8] Jia Yu and Rajkumar Buyya. Workflow Schdeduling Algorithms for Grid Computing Grid Computing and Distributed Systems (GRIDS) Laboratory Department of Computer Science and Software Engineering The University of Melbourne, Australia.

[9] Guangchang Ye, Ruonan Rao, Minglu Li. A Multiobjective Resources Scheduling Approach Based on Genetic Algorithms in Grid Environment. In Fifth International Conference on Grid and Cooperative Computing Workshops (GCCW'06) IEEE computer society.

[10] Taras S. Shapovalov, Alexey G. Tarasov. Genetic Algorithm Based Parallel Jobs Scheduling. In program "Research and scientific-pedagogical personnel of innovative Russia"(project No. 02-740-11-0626) and Grant of Russian Foundation for Basic Research and Far eastern branch of Russian academy of sciences No. 10-III-B- 01I-009.

[11] Wei Sun , Yuanyuan Zhang , Yanwei Wu, and Yasushi Inoguchi Practical Task Flow Scheduling for High Throughput Computational Grid. In International Conference on Parallel Processing Workshops (ICPPW'06) 0-7695-2637-3/06 $20.00 © 2006 IEEE computer society.

[12] T. Casavant, and J. Kuhl, A Taxonomy of Scheduling in General-purpose Distributed Computing Systems, in IEEE Trans. on Software Engineering Vol. 14, No.2, pp. 141--154, February 1988.

[13] Arash Ghorbannia Delavar, Mohsen Nejadkheirallah, Mehdi Motalleb. A New Scheduling Algorithm for Dynamic Task and Fault Tolerant in Heterogeneous Grid Systems Using Genetic Algorithm. In IEEE computer society 2010.

[14] Jing Liu, Li Chen, Yuqing Dun, Lingmin Liu, Ganggang Dong. The Research of Ant Colony and Genetic Algorithm in Grid Task Scheduling. In International Conference on MultiMedia and Information Technology 2008.

[15] S. Prabhu, V.Naveen Kumar. Multi-Objective Optimization Based on Genetic Algorithm in Grid Scheduling. International Journal of advanced research in technology. IJART, Vol. 1 Issue 1, 2011. ISSN NO: 6602 3127 RR.

[16] Dr. K.Vivekanandan, D.Ramyachitra A Study on Scheduling in Grid Environment Dr. K. Vivekanandan et al. / International Journal on Computer Science and Engineering (IJCSE).

Design Issues in Cloud-Hosted Applications

A. Seetha Ram Nagesh[1] and Suhail Afroz[2]
[1] CVR College of Engineering, Department of IT, Ibrahimpatan, R.R.District, A.P., India
Email: asnagesh@rediffmail.com
[2] CVR College of Engineering, Department of CSE, Ibrahimpatan, R.R.District, A.P., India Email:
suhailafroz@hotmail.com

Abstract – Cloud Computing is the hottest technology in the market these days, used to make storage of huge amounts of data and information easier for organizations. Maintaining servers to store all the information is quite expensive for individual and organizations. Cloud computing allows to store and maintain data on remote servers that are managed by Cloud Service Providers (CSP) .The concept of building or consuming services and applications that are hosted off-premises is becoming more attractive both to independent software vendors (ISVs) and to enterprises as a way to reduce costs, maximize efficiency, and extend capabilities. This paper describes the nature and use of cloud-hosted services and applications. It describes the benefits and the typical design issues, and the constraints and technology considerations often encountered when building and consuming these kinds of applications.

I. Introduction

From initial concept building to current actual deployment, cloud computing is growing more and more mature. Nowadays many organizations, especially Small and Medium Business (SMB) enterprises, are increasingly realizing the benefits by putting their applications and data into the cloud. The adoption of cloud computing may lead to gains in efficiency and effectiveness in developing and deployment and save the cost in purchasing and maintaining the infrastructure

Regarding definition of cloud computing model, the most widely used one is made by NIST as "Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models."[1] The cloud computing model NIST defined has three service models and four deployment models. The three service models, also called SPI model, are: Cloud Software as a Service (SaaS), Cloud Platform as a Service (PaaS) and Cloud Infrastructure as a Service (IaaS). The four deployment models are: Private cloud, Community cloud, Public cloud and Hybrid cloud.

Cloud computing represents the converging evolution of computing infrastructure and application models for building and consuming scalable distributed solutions. As techniques for building these kinds of applications have advanced, so too have the capabilities of the infrastructure on which they run. This synergistic evolution allows the infrastructure to be provisioned and maintained largely independently of the applications that it hosts. This in turn allows applications to take advantage of supporting infrastructure services and capabilities while they focus on their specific business functionality.

Many organizations have been able to realize the joint benefits of scalable application models and supporting infrastructure internally on-premises in their own data centres. However, it is the ability to leverage an off-premises out sourced application hosting infrastructure that is behind much of the excitement around cloud computing. The infrastructure provider focuses on hardware, networking, power, cooling, and the operating environment that supports application manageability, reliability, and scalability; leaving the organization free to focus on their application's business functionality. This provides many benefits in terms of reduced capital outlay and operating costs; and increased capacity, scalability and availability.

To leverage these benefits, cloud-hosted applications typically must be architected to follow a specific application model. This allows the cloud-hosting provider to generalize and optimize their operating environment support for application manageability, reliability, or scalability.

Different cloud-hosting providers have different application model requirements. Some adopt a virtual machine approach, where the application is developed and packaged along with its operating system image and the dependent runtime frameworks. Others utilize an application model that provides higher level abstractions for data access and storage , and for computation and communication. Still others provide higher level application models based on highly configurable applications that focus on specific vertical application functionality, such as Enterprise Resource Planning (ERP) or Customer Relationship Management (CRM). Each of these approaches provides distinct advantages and disadvantages.

Furthermore, some off-premises hosted applications are self-contained and designed for users who interact with the application through a dedicated UI. Some of these applications are service-enabled, and provide both a UI and expose their functionality through an API (often exposed through standards such as REST or SOAP) so that they can be integrated into other applications, which themselves can be hosted either on-premises or off-premises. Some off-premises hosted services are specifically designed to provide functionality for integration into other applications, and provide no UI at all.

The paper is organized as follows. Section II discusses the services provided by cloud environment. Section III explains the benefits of the Cloud applications. Section IV exploits the design issues for both the ISVs and Enterprise customers in developing Cloud-hosted applications. Section V summarizes the relevant patterns for design issues discussed in Section IV.

II. CLOUD SERVICES

Cloud-based services generally fall into categories such as storage/compute, business services, and retail/wholesale services. Some common examples of these remote services are:

- Business services such as stocks and shares information, invoicing and payment systems, data interchange facilities, merchant services, and business information portals.
- Retail/wholesale services such as catalogues, stock query and ordering systems, weather and traffic information, mapping services, and shopping portals.
- Storage/compute services such as data storage and processing, data backup, source control systems, and technical or scientific processing services.

These remote services can be consumed by software that runs on-premises, in an organization's data center or on a user's machine, which may be a desktop computer or any other Internet-enabled device. This typically involves a mix of technologies and techniques that are referred to as Software plus Services (S+S) [2]. S+S refers to an approach to application development that combines hosted services with locally executed software. The combination of the remote services and the software running locally, with rich seamlessly integrated interfaces and user experience, can provide a more comprehensive and efficient solution than traditional on-premises silo applications. S+S is an evolution of several other technologies including Service Oriented Architecture (SOA), Software as a Service (SaaS), Platform as a Service (PaaS), and Web 2.0 community-oriented architectural approaches.

A. Cloud-hosted Services

- Building block service: A service designed to be consumed by or integrated with other applications or services. An example is a storage service or a hosted Security Token Service (STS) such as the Access Control Service in the Azure Services Platform.
- Cloud-hosting environment: An environment that provides a core runtime for hosting applications; and, optionally, building block services, business services, social network services, and hosting services such as metering, billing, and management.
- Home-built application: An application that you create in-house, usually specifically targeted at some task, scenario, or process you require; it will often address a need that cannot be sourced from a third party.
- Hosted application: An application (packaged or home-built) hosted as a service. It may be hosted internally on your own system, or hosted externally by a partner or hoster.
- Packaged application: An application created by a third party or vendor that may provide only limited customization capabilities based on configuration or plug-ins.
- Platform as a Service (PaaS): A core hosting operating system, and optional plug-in building block services, that allow you to run your own applications or third-party applications obtained from vendors, in a remote cloud hosting environment.
- Software as a Service (SaaS): Applications that perform comprehensive business tasks, or accomplish business services, and allow you to consume them as services with no internal application requirements other than composition and UI.

III. BENEFITS OF CLOUD APPLICATIONS

Cloud-hosted applications and services may be very beneficial to ISVs, and to service delivery or hosting companies that build, host and deliver services. They also offer benefits to large enterprises that generally consume hosted and cloud-based solutions.

A. Benefits for ISVs and Service Hosts

The key advantages for ISVs and service hosting companies building and offering cloud-based solutions are the following:

- Architectural Flexibility: Vendors can offer their customers a range of deployment options, including hosting for the services they require, and allow users to choose from a range of prebuilt features or choose which features of the application they will implement themselves. This can reduce the architectural liabilities for end users who are developing services
- Rich User Experience: ISVs and service providers can offer richer experiences to their customers by leveraging existing specialized services (such as Virtual Earth). Hosters can combine their offerings with other cloud services obtained elsewhere to offer additional value propositions, and make it easier for end users to integrate services.
- Ubiquitous Access: Services in the cloud persist user data and state, and resynchronize when the user reconnects from any location. This supports both offline and occasionally connected scenarios, which is especially useful for mobile devices where a constant connection or bandwidth cannot be guaranteed.

ISVs and service hosts may also consider entering the market for commercial reasons to take advantage of monetization opportunities. The following are some examples:

- Vendors may wish to take advantage of an untapped market opportunity by offering a product that is not currently or easily available elsewhere, or use the cloud to offer lower end versions of their products to protect a high end franchise.
- Startup companies may use the cloud-hosted approach to minimize initial capital expenditure, and to take advantage of properties of the cloud such as elasticity (the capability to grow as required without high initial cost commitment).
- Vendors and users can create applications that generate income more quickly by taking advantage of ancillary services that are already available. For example, they can take advantage of payment and accounting systems in the cloud. Users can even build virtual stores without requiring large investments in IT equipment and networking capabilities.

B. Benefits for Enterprise Service Consumers

The key advantages for enterprises that consume cloud-based solutions are the following:

- Architectural Flexibility: In-house developers can create complete solutions that compose services in the cloud with local application code and their own services. IT departments can choose which features of the application they will implement themselves, and buy in other services that they require.
- Cost and Time Savings: IT departments can select the best cloud-based service for each task, and combine them to expose fully functional applications with shorter development times, and at a reduced cost. In addition, the reduction in the requirements for in-house IT infrastructure simplifies management, security, and maintenance costs.
- Economies of Scale: Companies can leverage economies of scale for industry average capabilities, and focus on their core activities. The economies of scale available from hosted applications arise from a range of factors, including reduced in-house infrastructure costs to better utilization of hardware that offers opportunities for reduced running costs. However, the gains in economies of scale must be balanced with the loss of control inherent with moving from on-premises to fully hosted applications.
- Offline Capability: The cloud can act as hub for roaming users. User data and state can be stored in the cloud and resynchronized when the user reconnects. Users can move between desktop and mobile clients seamlessly with fewer network configurations.

IV. Design Issues

Several common issues are of concern to both ISVs and enterprise customers [3]. While they cover a range of different aspects of hosted and cloud-based scenarios, these issues can be categorized into specific areas.

- Data Isolation and Sharing
- Data Security
- Data Storage and Extensibility
- Multi-tenancy
- Performance
- Service Composition
- Service Integration

a) Data Isolation and Sharing

Hosters can implement isolation and sharing for databases and for database schemas. There are three basic models:

- Separate Databases: Each tenant has a separate database containing their own data schemas. This has the advantage of being easy to implement, but the number of tenants per database server might be relatively low, with subsequent loss of efficiency, and the infrastructure cost of providing services can rise rapidly. It is most useful when tenants have specific data isolation or security requirements for which you can charge a supplement.
- Shared Databases, Separate Schemas: All tenants use the same database, but have separate sets of predefined fields available. This approach is also easy to implement, maximizes the number of tenants per database server, and improves database efficiency. However, it usually results in sparsely populated tables in the database. It is most useful when storing data for different tenants in the same tables (commingling) is acceptable in terms of security and isolation, and when you can anticipate the predefined custom fields that will be required.
- Shared Databases, Shared Schema: All tenants use the same database and special techniques are used to store data extensions. This approach has the advantage that the number of custom fields you can offer is practically unlimited. However, indexing, searching, querying, and updating processes are more complex. It is most useful when storing data for different tenants in the same tables (commingling) is acceptable in terms of security and isolation but it is difficult to predict the range of predefined custom fields that will be required.

b) Data Security

Cloud-hosted applications must implement strong security, using multiple defense levels that complement one another to provide data protection in different ways, under different circumstances, and against both internal and external threats [7]. When planning a security strategy, consider the following guidelines:

- Filtering: Use an intermediate layer between a tenant and a data source that acts as a sieve so that it appears to the tenant that theirs is the only data in the database. This is especially important if you use a shared database instance for all of your tenants.
- Permissions: Use access control lists (ACLs) to determine who can access data in the application, and what they can do with it.
- Encryption: Obscure every tenant's critical data so that it will remain unreadable to unauthorized parties, even if they manage to access it.

c) Data Security Patterns

Depending on the multi-tenant model adpoted, consider the following security patterns:

- Trusted Database Connections (applies to all three multi-tenant models): The application always connects to the database using its own application process identity, independent of the identity of the user, and the server grants the application access to the database objects that it can read or manipulate. Additional security must be implemented within the application itself to prevent individual end users from accessing any database objects that should not be exposed to them. Each tenant (organization) that uses the application has multiple sets of credentials associated with their tenant account, and must grant their end users access to the application using these credentials. These end users access the application using their individual credentials associated with the tenant account, but the application accesses the database using the single set of credentials associated with that application. This means that a single database access account is required for each application (one for each tenant). Alternatively, you can use an STS to obtain encrypted login credentials for the tenant irrespective of the individual user, and use security code in the application to control which data individual users can access.
- Secure Database Tables (applies to the Separate Database model and the Shared Database, Separate Schema model): Grant a tenant user account access to a table or other database object. In the Separate Database model, restrict access on a database-wide level to the tenant associated with that database. In the Shared Database, Separate Schema model, restrict access on a per table basis to the tenant associated with specific tables.
- Tenant Data Encryption (applies to all three multi-tenant models): Secure the data using symmetric encryption to protect it, and secure the tenant's private key using asymmetric (public/private key pair) encryption. Use impersonation to access the database using the tenant's security context, and use the tenant's private key to decrypt the data in the database so that it can be used. The disadvantage is that you cannot index encrypted columns, which means that there is a tradeoff between data security and performance. Try to avoid using index fields that contain sensitive data.
- Tenant Data Filter (applies to the Shared Database\Shared Schema model): Use SQL views to select subsets of data from tables based on the tenant or user ID, or the tenant account's security identifier. Grant tenants access to only their views, and not to the underlying tables. This prevents users from seeing or accessing any rows belonging to other tenants or users in the shared tables.

d) Data Storage and Extensibility

Hosted data may be stored in variety of ways. Two different approaches are emerging for implementing data storage in hosted applications: hosted relational database management systems (RDBMS) and non-relational cloud-based storage. Relational database systems provide storage for structured data, and are more suited to transactional systems or applications that are I/O intensive; they also typically provide lower latency and advanced query capabilities. In contrast, cloud storage refers to any type of data storage that resides in the cloud; including services that provide database-like functionality, unstructured data services (for example, file storage for digital media), data synchronization services, and network-attached storage (NAS) services. Data services are often consumed in a pay as you go model, or in this case a pay per GB model (including both stored and transferred data).

Cloud storage offers a number of benefits, such as the ability to store and retrieve large amounts of data in any location at any time. Data storage services are fast, inexpensive, and almost infinitely scalable; however, reliability can be an issue as even the best services do sometimes fail [6]. Applications that are sensitive to high latency might also be affected as each interaction with the storage service requires network transversal. Finally, transaction support can be an issue with cloud-based storage systems. These systems generally focus heavily on partitioning and availability, and consistency cannot always be guaranteed.

e) Multi-tenancy

The idea of multi-tenancy, or many tenants sharing resources, is fundamental to cloud computing. Service providers are able to build network infrastructures and data architectures that are computationally very efficient, highly scalable, and easily incremented to serve the many customers that share them. Multi-tenancy spans the layers at which services are provided [9]. In IaaS, tenants share infrastructure resources like hardware, computer servers, and data storage devices. With SaaS, tenants are sourcing the same application (e.g., Salesforce.com), which means that data of multiple tenants is likely stored in the same database and may even share the same tables. When it comes to security, the risks with multi-tenancy must be addressed at all layers.

Individual tenants share the use of the hoster's hardware and infrastructure, as well as sharing databases and database systems. Service suppliers must provide a platform with appropriate capacity and performance for hosted services. They must also consider how to keep the cost structure under control, and how they will provide customization through configuration. There are four common stages in moving towards an efficient multi-tenancy architecture with user-enabled configuration. The following sections describe these stages.

- Custom: Each customer runs a separate copy of the software assigned only to that customer, and the only way to support multiple customers is to serve them with different copies of the software. Furthermore, because little is done to allow customization through configuration, each copy includes specific customer customizations in the form of custom extension code, custom processes, and/or custom data extensions. Although the software is, technically, delivered as a service (it does not run on the customer's premises), economy of scale cannot be achieved because each customer runs a different instance of the software. Although this could be a useful starting point to validate the business model, it must be avoided once the volume of customers increases. It is impractical to manage thousands of customers using this model.
- Configurable: The software can be tailored for each tenant through configuration and by avoiding the use of custom code. All the tenants run the same code; however, the architecture is still not multi-tenant and each customer runs their own copy of the code, even though the copies are identical. The separation can be either virtual (virtual machines on a same server) or physical (running on separate machines). Although this model is a considerable improvement over the custom model described above, the architecture still allows customization through configuration, and the computing power is not shared among the instances. Therefore, the provider cannot achieve economy of scale.
- Multi-tenant: The UI can be customizable per tenant, as can the business rules and the data model. The customization per tenant is entirely through configuration using a self service tool, which removes the requirement for the service provider to perform configuration. This level is almost the SaaS perfect case; the exception is any capacity to scale out. At this level, data partitioning means that growth can only be achieved by scaling up.
- Scalable: The architecture supports multi-tenancy and configuration, plus the capability to scale out the application. New instances of the software can be transparently added to the instance pool to dynamically support the increasing load. Appropriate data partitioning, stateless component design, and shared metadata access are part of the design. At this level, a Tenant Load Balancer (implemented using a round robin or a rule based mechanism) is introduced, maximizing the utilization of hosting resources such as CPU and storage.

This means that the total load is distributed across the entire available infrastructure. The data is also reorganized periodically in order to average the data load per instance. The architecture is scalable, multi-tenant, and customizable through configuration.

f) Performance

Cloud-hosted applications must be scalable to support increasing numbers of services, and increasing load for each service and tenant [11]. When designing services, consider the following guidelines for scaling applications:

- Design services and components to be stateless where possible. This minimizes memory usage for the service, and improves the opportunity to scale out and load balance servers.
- Use asynchronous input and output calls, which allow the applications to do useful work while waiting for I/O to complete.
- Investigate the capabilities of the hosting platform that can improve performance. For example, in Microsoft Azure, use queues to manage requests and worker processes to carry out background processing.
- Use resource pooling for threads, network, and database connections.
- Maximize concurrency by using locking only where absolutely necessary.

When scaling data storage and applications, consider the following guidelines:

- When scaling the data partition, divide subscriber data into smaller partitions to meet performance goals. Use schemes such as Hashing (to subdivide content) and Temporal (based on the time or date range in which the data is valid).
- Consider implementing dynamic repartitioning to repartition the data automatically when the database size reaches a specific maximum size.

When scaling data storage and applications investigate standard patterns, and the specific techniques and implementations provided by the hosting platform—some examples are data partitioning, load balancing, failover, and geographical distribution.

g) Service Composition

Users in enterprise-level organizations require access to many different document repositories, types of data, sources of information, and applications that perform specific functions. Traditionally, users interacted directly with each store or application, often using specific isolated applications. However, over time, enterprises have attempted to consolidate systems; often using intranet Web portals or façade-style applications that connect to the appropriate downstream applications.

With the advent of services and SOA applications, IT departments can expose applications and data as services, either hosted in-house or bought in as SaaS. The service portfolios can still expose the combination of traditional local applications, internally hosted services, and remote services through portals, which hide the user from the implementations and allow IT departments to adapt the ranges of services quickly and easily. However, S+S and SaaS designs and technologies allow IT departments and enterprise customers to integrate services fully. Service integration can help to achieve the goal of a *many to one* model where all applications and services are available to the user through a composition architecture that effectively exposes them as a single application, as shown in Figure 1. A service integration mechanism combines the groups of applications in the portfolios and exposes them though a rich client that can interact with any service or application.

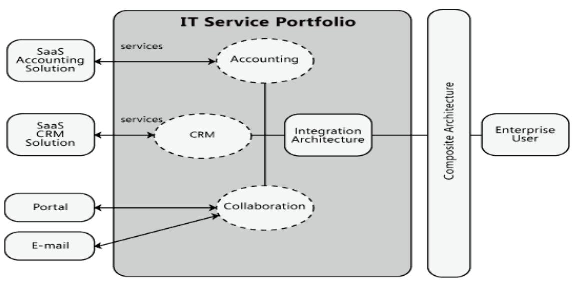


Figure1. Service Integration

h) Service Integration

Cloud-hosted solutions can help to mitigate some of the challenges encountered with traditional software, but add new and different challenges for the consumer of these services. Consider the following the challenges when moving to hosted cloud services and applications:

- Identity Management: Enterprise procedures for adding, updating, and removing users must be extended to the remote services. If the external service depends on user identity, which is very likely for SaaS and for S+S, the provisioning and deprovisioning processes must be extended. In addition, translation of in-house user identity into specific roles may be required, possibly through a federated service, to minimize the migration or duplication of individual user identities at the remote service host. Enterprise user account policies such as password complexity and account lockouts must also be compatible with those of the remote service supplier. If no SSO facility is available, there can be increased liabilities, maintenance costs, and operational inefficiencies.
- Data: Requirements of data operations, such as Extract, Transform, and Load and data integration, must be analyzed for compatibility with service capabilities. Hosted services may not support complex data storage patterns, which may affect the design of data entities and application architecture. In addition, data may need to be protected more securely to counterbalance the lack of physical security available when hosting in-house. However, applications can store sensitive or private data locally, and use the cloud services only for nonsensitive data.
- Operations: In-house integration services and client applications may not be compatible with services exposed by the service supplier, even when using industry standard protocols. You must also ensure that the service provider can generate appropriate reporting information, and determine how you will integrate this with your own management and reporting systems. In terms of service levels, Service Level Agreements (SLAs) may require revision to ensure that they can still be met when depending on the service provider for escalated support. Enterprises must also be prepared to implement help desk facilities that act as the first contact point for users, and define procedures for escalating issues with the service provider.
- Security: Enterprise privacy policies must be compatible with those of the service provider, and rules for actions that users can execute, such as limits on transaction size and other business rules, must be maintained—even if these are not part of the remote service capabilities. This may make the service integration infrastructure more complex. Procedures and policies for maintaining the security and integrity of data in the event of service or interconnectivity failure will also be required. Authentication, encryption, and the use of digital signatures will require the purchase of certificates from certified providers, and may require implementation of a Public Key Infrastructure (PKI). In addition, integration may require changes to firewall rules, and updates to firewall hardware and software may need to be required to provide filtering for application data and XML Schema validation.
- Connectivity: Some types of cloud-based applications rely on good quality broadband Internet connections to function well. Examples are online transaction processing and real time services such as voice over IP (VoIP) and Microsoft Office Communications Server. In some areas and some countries, this may not be available. In addition, services that require large data transfers such as backup services and file delivery services will generally run more slowly over an Internet connection compared to a local or in-house implementation, which may be an issue. However, messaging and other similar services may not be as dependent on connection

bandwidth or severely affected by occasional loss of connectivity.

- Service Level Agreements: Skills and expertise will be required to assess suppliers more comprehensively, and make choices regarding service acquisition and contracts. SLAs may also require revision to ensure that they can still be met when depending on the services hosted by a remote provider.
- Compliance and Legal Obligations: Compliance with legal and corporate directives may be affected by the performance of the service supplier, or these compliance directives and legal obligations may conflict if the service provider is located in another country or region. There may also be costs associated with obtaining compliance reports from the service supplier. Local laws and policies may prevent some types of applications, such as banking applications, from running in hosted scenarios.

V RELEVANT DESIGN PATTERNS

Key patterns are organized into categories such as Data Availability, Data Transfer, Data Transformation, Integration and Composition, Performance and Reliability, and User Experience as shown in the following table. Consider using these patterns when making design decisions for each category.

Category	Relevant patterns
Data Availability	Polling: One source queries the other for changes, typically at regular intervals. Push: A source with changed data communicates changes to the data sink every time data in a data source changes, or only at regular intervals. Publish/Subscribe: A hybrid approach that combines aspects of both polling and pushing. When a change is made to a data source, it publishes a change notification event, to which the data sink can subscribe.
Data Transfer	Asynchronous Data Transfer: A message-based method where the sender and receiver exchange data without waiting for a response. Synchronous Data Transfer. An interface-based method where the sender and receiver exchange data in real time.
Data Transformation	Shared Database: All applications that you are integrating read data directly from the same database. Maintain Data Copies: Maintain copies of the application's database so that other applications can read the data (and potentially update it). File Transfer: Make the data available by transporting a file that is an extract from the application's database so that other applications can load the data from the files.
Integration and Composition	Broker: Hide the implementation details of remote service invocation by encapsulating them into a layer other than the business component itself. Composition: Combine multiple services, applications, or documents into an integrated interface while performing security, validation, transformation, and related tasks on each data source. Portal Integration: Create a portal application that displays the information retrieved from multiple applications within a unified UI. The user can then perform the required tasks based on the information displayed in this portal.
Performance and Reliability	Server Clustering: Design your application infrastructure so that your servers appear to users and applications as virtual unified computing resources to enhance availability, scalability, or both. Load-Balanced Cluster. Install your service or application onto multiple servers that are configured to share the workload. The load-balanced hosts concurrently respond to different client requests, even multiple requests from the same client. Failover Cluster. Install your application or service on multiple servers that are configured to take over for one another when a failure occurs. Each server in the cluster has at least one other server in the cluster identified as its standby server.

Table I.
Patterns to be considered while Designing Cloud-Hosted Applications

CONCLUSION

Cloud Computing is the cost, time and performance effective. Some basic issues are the key concern in the Cloud Computing use and in the implementation for the Client as well as for Vendors.Cloud computing allows to store and maintain data on remote servers that are managed by Cloud Service Providers (CSP) .The concept of building or consuming services and applications that are hosted off-premises is becoming more attractive both to independent software vendors (ISVs) and to enterprises as a way to reduce costs, maximize efficiency, and extend capabilities.The current technology does not provide all the requirements needed by the cloud computing. There are many challenges to be addressed by the researchers for making cloud computing work well in reality. Some of the challenges like security issues and Data issues are very much required for the customers to use the services provided by the cloud. Similarly challenges like Security, performance issues and other issues like service comspostion etc are important for the service providers to improve the services. In this paper we have identified the challenges in terms of security issues, data challenges, performance challenges and other design challenges. We have provided an insight into the possible solutions to these problems even though lot of work is needed to be done in this regard

REFERENCES

[1] Peter Mell, and Tim Grance,“The NIST Definition of Cloud Computing,” Version15, 10-7-09 http://www.wheresmyserver.co.nz/storage/media/faq-files/ cloud-def-v15.pdf.

[2] "Software + Services (S+S)" at http://msdn.microsoft.com/en-us/architecture/aa699384.aspx

[3] Traian Andrei, ”Cloud Computing Challenges and Related Security Issues”, Survey Paper

[4] Kresimir Popovic, et al., “Cloud Computing issues and challanges” MIPRO 2010 May 24-28 Opatija, Croatia, pages 344-349.

[5] Shucheng Yu, Cong Wang, Kui Ren, and Wenjing Lou, "Achieving Secure, Scalable, and Fine-grained Data Dccess Control in Cloud Computing", The 29th IEEE Conference on Computer Communications (INFOCOM'10), San Diego, CA, March 15-19, 2010.

[6] Bhavani Thuraisingham, Vaibhav Khadilkar, Anuj Gupta, Murat Kantarcioglu, Latifur Khan, “Secure Data Storage and Retrieval in the Cloud”, The University of Texas at Dallas

[7] L. Kaufman. “Data security in the world of cloud computing”. IEEE SECURITY & PRIVACY, 7(4), July- August 2009.

[8] Kevin Hamlen, Murat Kantarcioglu, Latifur Khan, Bhavani Thuraisingham, “Security Issues for Cloud Computing”, International Journal of Information Security and Privacy, 4(2), April-June 2010

[9] "Multi-Tenant Data Architecture" at http://msdn.microsoft.com/en-us/architecture/aa479086.aspx

[10] Sadie Creese, Paul Hopkins, Siani Pearson, Yun Shen, “Data Protection-Aware Design for Cloud Computing”,HPL-2009-192

[11] Michael Olson, K. Mani Chandy,“Performance issues in cloud computing for cyber-physical applications”

Architectural Design Patterns Customized and Validated for Flight Software

U.V.R. Sarma[1], N. Pavani[2] and Dr. P. Premchand[3]
[1] CVR College of Engineering, Department of CSE, Ibrahimpatan, R.R. District, A.P., India
Email: sarmauvr@yahoo.co.in
[2] CVR College of Engineering, Department of CSE, Ibrahimpatan, R.R. District, A.P., India
Email: mee_pav@yahoo.com
[3]Osmania University, Department of CSE, Hyderabad, A.P., India
Email: p.premchand@uceou.edu

***Abstract*— Software design patterns are best practice solutions to common software design problems. When they are properly applied, software design patterns can greatly improve the quality of software architectures. Leveraging the benefits of design patterns is particularly important in the space Flight Software (FSW) domain because better designs are needed to help reduce the number of flight software related anomalies and thus improve the quality of FSW architectures. This paper provides a solution to build templates for common features of Flight Software architecture using design patterns. This is illustrated by using Student Nitric Oxide Explorer (SNOE) spacecraft, which is a real world case study from National Aeronautics and Space Administration (NASA). The executable design pattern templates help an engineer when building software architectures. This paper also provides a foundation to perform validation for functional correctness during the design phase.**

***Key Words:* Software Architectural Design Patterns, UML, Space Flight Software, IBM Rational Rhapsody.**

I. Introduction

Software design patterns are best practice solutions to common software problems. Design patterns are normally captured to be domain and platform independent. There are several benefits of capturing design patterns in this manner. First, it makes them applicable across multiple domains and platforms. Second, it makes design patterns applicable at different levels of abstraction. Furthermore, in the majority of cases, multiple design patterns can be applied in a single application.

To achieve this goal, this paper provides a set of design patterns that are applicable to a small satellite Student Nitric Oxide Explorer (SNOE). This paper also describes a validation approach that is used to validate the functionality of software architectures.

This paper is applied and validated using the space Flight Software (FSW) domain. FSW is an ideal domain to apply this dissertation for multiple reasons. First, the amount of requirements and responsibilities placed on FSW is growing. FSW has evolved from performing simple operations to controlling a majority of the spacecraft payloads. This paper is a way to architect FSW using design patterns. Using design patterns makes certain that best practices are incorporated into FSW designs.

Secondly, the industry trend indicates that the number of software related anomalies is growing. It is reported that “in the period from 1998 to 2000, nearly half of all observed spacecraft anomalies were related to software” [1]. These software anomalies can cause mission disruption or even mission loss. In the aerospace industry these losses cannot be tolerated because of the high cost and length of time that is required to build a spacecraft. Additionally, many spacecrafts support very critical missions that can be severely impacted from a small disruption of service. This paper helps to alleviate the number of software related anomalies by providing design time validation. Therefore, design flaws that lead to software anomalies can be identified and remedied early.

This paper is organized as follows. First it describes about IBM Rational Rhapsody. Next, about UML 2.0 and how it was used in the paper, then about SNOE and the process for customizing general design patterns for SNOE using IBM Rational Rhapsody is described in detail. Finally, this paper includes a discussion on conclusions and areas of future work.

II. IBM Rational Rhapsody

This project uses IBM Rational Rhapsody to build and execute the state machines [2]. Therefore the actions performed are captured using IBM Rational Rhapsody’s action language and event handling infrastructure. IBM Rational Rhapsody uses custom action language, which is a subset of the Java language, to capture actions and to execute the model. Thus, this action language is used to implement the objects actions. The action language is similar to Java, except there are a few additional reserved words and functions. For example, GEN is a reserved word used to generate asynchronous messages as events. The messages must be specified on the consumer’s provided interface in order to be invoked.

Ex: PClass1.gen(new msg());

Where PClass1 is the provided interface which also specifies the port through which the message is sent and msg() is event that is generated. When an event is generated, IBM Rational Rhapsody event handling infrastructure handles the routing of events from the producer to the consumer. When the consumer component receives the event, the appropriate state transition is taken and actions within that state are performed. IBM Rational Rhapsody is an excellent tool to generate dynamic UML diagrams using Real-time UML that is UML 2.0.

III. UML 2.0 As Architectural Description Language (ADL)

The Unified Modeling Language (UML) [3] is formal graphical language considered as a de facto industrial standard. Although the language has been created as graphical language firstly to support object oriented software analysis and design, the language has been revised couple of times and today, it is a general formal language capable to describe a software system. The UML has well defined formal syntax and semantics and can be machine checked and processed. UML includes a set of graphical notation techniques to create abstract models of specific systems.

The expressive power of Architectures by UML is more than any ADL. The UML profile for scheduling, performance, and time specification described in [4] has been adopted as an official OMG standard in March 2002. In general, UML profile defines a domain specific interpretation of UML; it might be viewed as a package of specializations of general UML concepts that capture domain-specific variations and usage patterns. To specify a profile, UML extensibility mechanisms (i.e., stereotypes, tagged values, constraints) are used.

Component and connector views (C&C views, for short) present an architecture in terms of elements that have a runtime presence (e.g., processes, clients, and data stores) and pathways of interaction (e.g., communication links and protocols, information flows, and access to shared resources). *Components* are the principal units of run-time interaction or data storage. C*onnectors* are the interaction mechanisms among components.

The components are created as Composite classes in UML 2.0 and each of the components should have ports to interact with the external environment. Each port again requires an interface for it to interact. The interfaces are of two types *Provided Interface* and *Required Interface.* Two components with ports and their interfaces can be linked for communication. The ports and their interfaces should be compatible, that is one component having a required interface (depicted as semi circle) can interact with only a component that provides the interface (depicted as full circle).

IV. Student Nitric Oxide Explorer (SNOE)

This paper illustrates the construction of architecture for Flight Software by taking up a case study of Student Nitric Oxide Explorer (SNOE) [5]. SNOE, which was a real-world, small satellite program funded by the National Aeronautics and Space Administration (NASA) and managed by the Universities Space Research Association (USRA).

SNOE's mission involves using a spin stabilized spacecraft in a low earth orbit to measure thermospheric Nitric Oxide (NO) and its variability. The SNOE spacecraft is spin stabilized, meaning it maintains its orientation similar to that of a top. SNOE is required to maintain a spin rate of 5 Rotations per Minute (RPM). The spin rate can be adjusted having the Flight Software (FSW) send a command to commutate the electromagnet transverse torque rod. The spin axis direction is controlled in a similar fashion by having the FSW send a command to commutate the electromagnet spin axis torque rod. SNOE's FSW does not perform the attitude determination and control calculations. Rather, the FSW collects the attitude measurements and downlinks them to the ground for processing.

Then the ground uplinks attitude control commands back to the spacecraft for the SNOE FSW to execute. The attitude measurements are taken from two Horizon Crossing Indicators (HCI) and three magnetometers. SNOE's spacecraft body is surrounded on all sides by stationary solar panels which are used to generate power.

The spacecraft contains three payload instruments to accomplish its scientific mission. These three instruments are an Ultra Violet Spectrometer (UVS) that measures NO density, an Auroral Photometer (AP) that measures the flux of energetic electrons entering the Earth's upper atmosphere, and a Solar soft X-ray Photometer (SXP) that measures the solar irradiance.

Additionally, SNOE also contains a microGPS Bit-Grabber Space Receiver (microGPS BGSR) instrument as a technology experiment. The microGPS BGSR gathers position information based on the Global Positioning System (GPS) constellation for experimental orbital determination.

4.1 SNOE Design Pattern Selection

The pattern selection process is done using the command execution functionality, which is a commonly seen in FSW. This involves determining the order in which spacecraft commands are executed. The design patterns that support this feature are then selected. For example, on small spacecraft the centralized control design pattern [6] can be used. The centralized control design pattern involves a single controller that provides overall control by conceptually executing a state machine. This design pattern is useful on small spacecraft because it encapsulates all the state-dependent control in a single component thus making the control logic easier to understand and maintain. Thus, the design patterns that support SNOE specific features are determined by selecting the Design Patterns that are suitable for the working of SNOE.

The paper illustrates the customization of Design Patterns to suit the architecture of the satellite Student Nitric Oxide Explorer (SNOE). Seven different Design Patterns have been identified to reflect the functionality of SNOE.

The Design Patterns identified are listed in Table I.

Table I. SNOE Design Pattern Selection

Feature	Design Pattern
Command Execution	Centralized Control Design Pattern
Telemetry Storage and Retrieval	Telemetry Client Server Design Pattern
Telemetry Formation	Pipes and Filters Design Pattern
Ground Driven Payload Data Collection	Payload Multiple Client Multiple Server Design Pattern
Ground Driven Housekeeping Data Collection	Housekeeping Multiple Client Multiple Server Design Pattern
Spacecraft Clock	Spacecraft Clock Multicast Design Pattern
Memory Storage Device Fault Detection	Memory Storage Device Watchdog Design Pattern

The reason for selecting the above Design patterns is described below:

4.1.1. Centralized Control Design Pattern: SNOE is a small satellite with thirteen different components. Since it is a small satellite and the number of components is less, Centralized control architecture is better suitable than Distributed architecture. The Centralized controller is linked to every component and controls the functionalities of each of the components.

4.1.2. Telemetry Client Server Design Pattern: The information collected by various components in SNOE is transformed into telemetry packets and is sent to the Ground Station. Every component has its own buffer and stores the information collected by them in their buffers. Next, the information is to be periodically transformed into telemetry packets and is to be sent to the Ground Station for processing. So a Client and Server component is created for each of the components which will be controlled by the Centralized Controller and will be responsible to collect the information. This pattern collects information from Payload Server as well as HouseKeeping Server and sends it to the controller.

4.1.3. Telemetry Formation Pipes and Filters Design Pattern: The transformation of information into telemetry packets is done by Pipes and Filters Design Pattern. It increases throughput capacity of the system by adding multiple homogeneous (identical) channels.

4.1.4. Payload Multiple Client Multiple Server Design Pattern: There are four payload instruments in SNOE. They are Ultra Violet Spectrometer, Micro GPS, Solar XRay Photometer and Auroral Photometer. A separate client and server for each of the payload instruments are created to collect the information whenever the controller signals to collect.

4.1.5. Housekeeping Multiple Client Multiple Server Design Pattern: The health of the satellite is maintained by collecting the information of the health or working of each of the component. This information is sent to the ground station. The ground station checks this information and sends any signals if necessary to check and modify the components. The collection of housekeeping information is done by this Design Pattern. Again a separate client and server component is created for 13 components of SNOE.

4.1.6. Spacecraft Clock Multicast Design Pattern: This pattern is used to send time signals to the Centralized controller and input and output components of the system.

4.1.7. Memory Storage Device Watchdog Design Pattern: The memory storage device in SNOE is EEPROM. The Memory Storage Watchdog Design Pattern is selected to check the working of the memory storage device that is the EEPROM at regular intervals.

V. Implementation

5.1 *SNOE Centralized Control Architectural Design Pattern*

SNOE utilizes the Centralized Control design pattern to execute commands and control the overall operation of the spacecraft. SNOE uses two torque rods, thus its multiplicity is one or many. Additionally, the ports and interfaces for the payload variants that are unique to SNOE are modeled. The component diagram for SNOE's Centralized Control component diagram is shown in fig. 2, which contains the SNOE specific variants based on SNOE's features. The ports, interfaces, and connecters for the common variants are captured in the diagram.

SNOE contains four payload devices - therefore four payload device variants are created. For each payload variant, the port name is updated to reflect the specific payload, such as the *microGPS_IOC*. The port's interface is updated to reflect the specific actions that can be invoked on that payload.

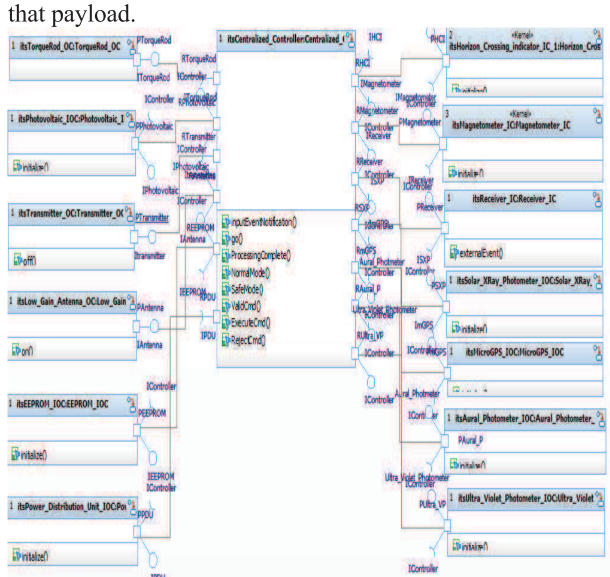


Figure 2. Component diagram for SNOE Centralized Control Executable Design Pattern

Next, the executable version of the design pattern involves potentially adding application specific states, actions, and activities to the state machines based on the application's features. For example, if the application features refine some behavior, then this can be modeled as sub-states. Also, if the component must send a message to an application specific variant or if application specific logic is required then this is modeled as an action or activity within a state or transition.

SNOE receives command to open the solar x-ray photometer door; it knows the precise operations to invoke on the Solar_Xray_Photometer_IOC. The state machine for the Solar_Xray_Photometer_IOC component is depicted in the fig. 3. The state machine for this Component is slightly more complicated because it acts as both an input and IO component. The component begins in the Idle state within the Working state. In the Idle state the Component waits for commands from the Centralized_Controller. When an action message is received, it transitions into the Executing_Command state where it performs the appropriate actions on the external hardware. After it performs the necessary actions, it generates the processingComplete event and transitions back to the Idle state to wait for the next command. When a read message is received, a similar set of states and transitions occurs, however, it occurs in the Gathering Data state. The IO_Component is also responsible for listening to external events from the hardware. Therefore if an externalEvent event is received, the IO_Component stops its current action in the Working state and transitions into the Preparing_Notification state. In the Preparing_Notification state it prepares a message to send to the Centralized_Controller.

Once the message is ready, the IO_Component then sends the inputEventNotification message to the Centralized_Controller through the RIO port and transitions back to its previously interrupted location within the Working state.

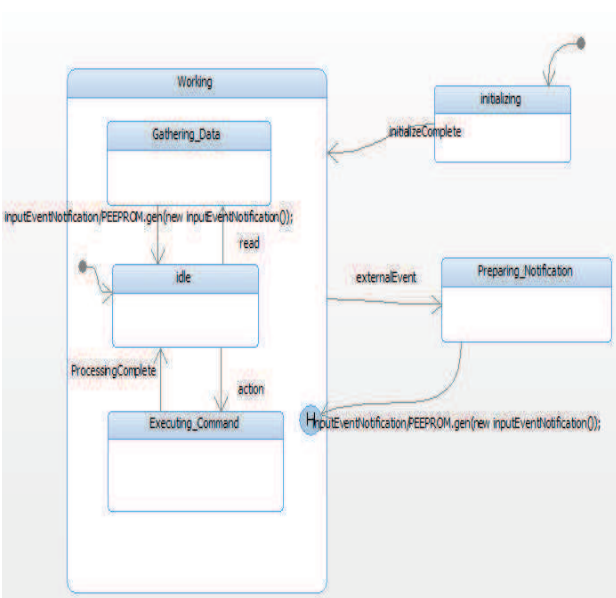


Figure 3. State Machine for Solar_XRay_Photometer_IOC

Next the state machine for the SNOE's Magnetometer_IC component is depicted in Figure 4. Magnetometer is an input component that provides attitude measurements. It is initialized by the Centralized Controller. It is first in the idle state and moves to the Preparing_Notification state when an external event occurs. Here it prepares the input_event_notification and sends it to the Centralized_Controller. A similar set of actions is performed in response to a read event message; however the requested data is collected and sent back the Centralized_Controller.

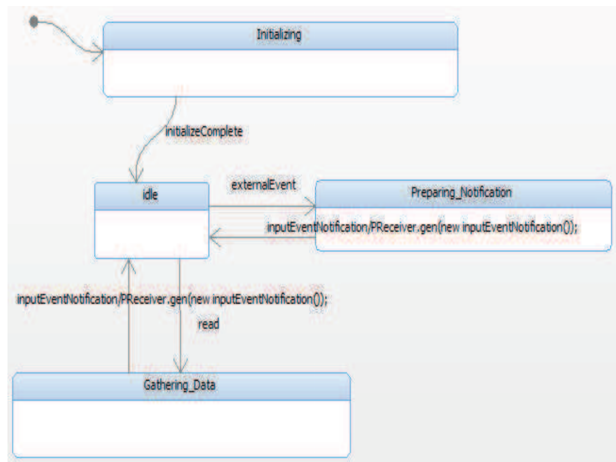


Figure 4. State Machine for Magnetometer_IC

The Output Component begins in the Idle state where is waits for commands from the Centralized_Controller. Once a command message is received, the Output_Component transitions into the Executing_Command state where it performs the appropriate actions on the external hardware. At the DRE level, the actions the Output_Component performs are variable; therefore it is modeled using a code stub as seen in fig. 5. Once complete, it generates the processing Complete event and transitions back to the Idle state to wait for the next command.

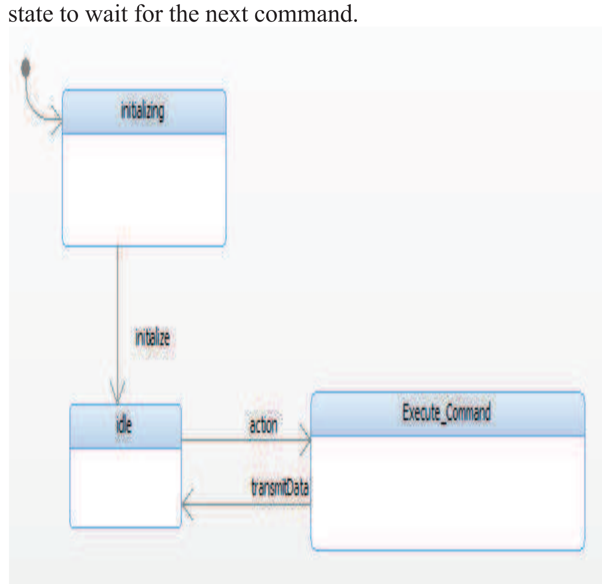


Figure 5. State Machine for Low_Gain_Antenna_OC

Finally, the state machines for the other variant input, output, and IO components, are also added.

5.2 SNOE Payload Multiple Client Multiple Server Executable Architectural Design Pattern

The next executable design pattern realized in SNOE is the FSW Payload Multiple Client Multiple Server executable design pattern. This design pattern is used to selectively collect payload data. Since SNOE is required to selectively collect the payload data, separate data clients are created for each payload instrument. Additionally, since each payload instrument has its own data buffer, separate server components are created for each payload instrument.

Next, the component diagram in fig. 6 depicts the set of components in the system. The ports and connectors added between the appropriate clients and servers are also shown in the diagram. Additionally, the interfaces are also updated to reflect the SNOE's unique variants. The diagram shows that the connected components have compatible interfaces.

The four Payload instruments include Ultra_Violet_Spectrometer, Micro_GPS, Solar_XRay_Photometer and Aural_Photometer. A client and server component for each of the four payload instruments is depicted in the design pattern.

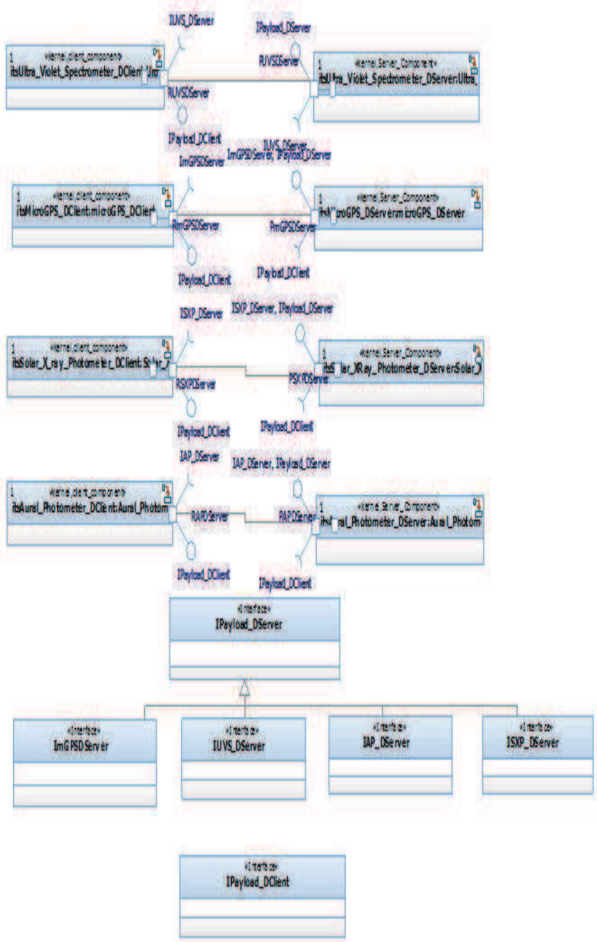


Figure 6. Object Model Diagram for Payload Multiple Client Multiple Server

The SNOE Multiple Client Multiple Server design pattern involves selectively collecting payload data. The interaction diagram for collecting micro GPS (Global Positioning System) data is depicted in Figure 7.

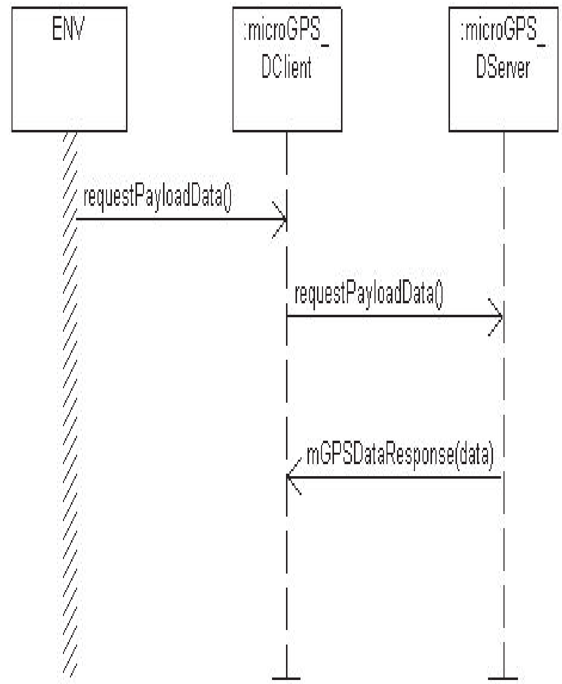


Figure 7. Collect microGPS Data Scenario for SNOE

5.3 SNOE Payload Multiple Client Multiple Server Executable Design Pattern

In addition to updating the architectural views, the executable version of the design pattern also needs to be customized for SNOE. This is performed for each client and server in this design pattern. The specific steps involved in updating the state machine are follows.

First, the microGPS_DClient component is responsible for collecting the microGPS data from the microGPS_DServer. The state machine for the SNOE specific microGPS_DClient component is depicted in fig. 8.

When Controller requires data it sends requestPayloadDataNeeded message to microGPS_DClient. microGPS_DClient requests the data from the server, this information is added to the actions on the state machine. This information is captured on the transition from the Preparing_Request state to the Idle state. The event that occurs is the requestPayloadData and the action

RUVSDServer.gen(newrequestPayloadData(msg));

Indicates that a request for payload data is being sent to the microGPS_DServer component by specifying the required port (RmGPSDServer) of the client through which the components communicate. Finally, the SNOE specific processing logic within the Preparing_Request and Processing_Response states is added as On Entry actions. However, this information is not depicted in Figure in an effort to make the diagram readable.

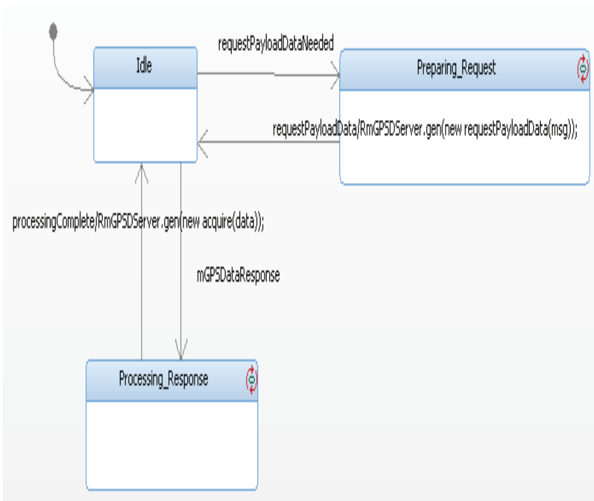


Figure 8, State Machine for MicroGPS Client

The state machine for microGPS_DServer in Figure 9. depicts the transitions that server takes. It is in Idle state first and moves to Processing_Client_Request state when client sends a requestPayloadData to server. After processing is complete, it prepares a response and moves back to the Idle state. During this transition it sends the mGPSDataResponse back to the client.

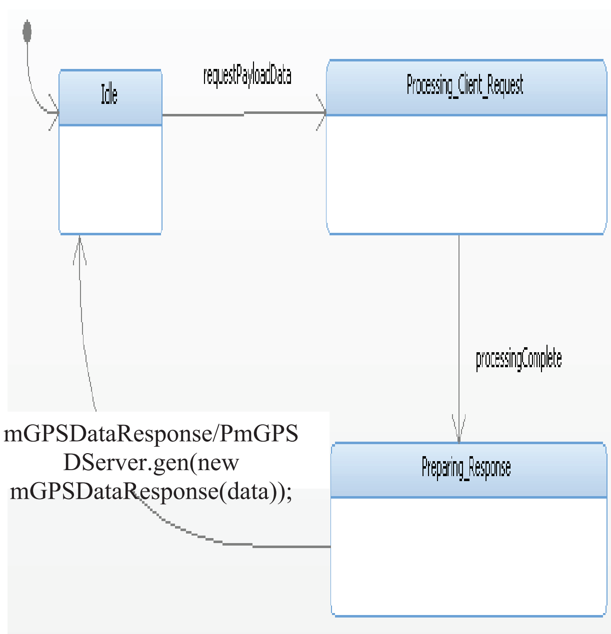


Figure 9. microGPS_DServer state machine

Similarly, the state machines for client and server for the other three payload instruments which are Auroral_Photometer, Solar_Xray_Photometer and Ultraviolet_Spectrometer are also updated following a similar process.

CONCLUSIONS

This paper describes an approach for building FSW software architectures from software architectural patterns. This approach improves the quality of FSW software architectures because it leverages best practices captured in software design patterns. Additionally, the executable design pattern templates not only help an engineer when building software architectures, but they also provide the foundation for performing design time validation on the software architecture produced using this approach. The engineers also can use the design patterns to form the core base for building the software architecture of any other system in this domain. Thus enabling to develop using the Software Product Line (SPL) based product development.

FUTURE ENHANCEMENTS

There are several avenues of future research that can be taken to extend this project. First, the SNOE case study can be expanded to include performance validation using MARTE (Modeling and Analysis of Real-Time Embedded systems) stereotypes. Second, this work should be extended to address feature modeling to help organize and structure the functionality of design patterns. Thirdly, this work can be applied to other DRE domains to illustrate the approach's applicability across DRE domains. Additionally, future research can include illustrating the functionality of the design patterns using the "animation" feature of IBM Rational Rhapsody.

REFERENCES

[1] Julie Street Fant, Hassan Gomaa, Robert G. Pettit, *Architectural Design Patterns for Flight Software*, 14th IEEE International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing Workshops, 2011.

[2] D. Harel, *Executable object modeling with statecharts*, *18th International Conference on Software Engineering*, 1997.

[3] B.Bharathi, Dr.D.Sridharan, *UML as an Architecture Description Language*, International Journal of Recent Trends in Engineering, Vol. 1, No. 2, May 2009

[4] Clements. P. et.al.: *Documenting Software Architectures, Views and Beyond*, Addison-Wesley, Boston, MA, USA,2002.

[5] Laboratory For Atmospheric and Space Physics at the University of Colorado at Boulder, *Student Nitric Oxide Explorer Homepage*, *http://lasp.colorado.edu/snoe/*. [Online]. Available: http://lasp.colorado.edu/snoe/. [Accessed: 21-Apr-2010].

[6] H. Gomaa, *Designing Software Product Lines with UML: From Use Cases to Pattern-Based Software Architectures*, Addison-Wesley Object Technology Series,, 2005.

Dual Band Monopole Antenna for Wireless Communications

A.Bharathi
CVR College of Engineering, Department of ECE, Ibrahimpatan, R.R.District, A.P., India
Email: bharathig8@gmail.com

Abstract—**Dual band antennas are used in many applications like in cellular phones, in laptop for wireless printer and modem connections. This paper discusses a dual monopole Microstrip antenna configuration for dual frequency coverage at 1800MHz for mobile communication and at 2400MHz for wireless LAN applications. The antenna consists of two monopoles of different stems printed on a FR4 substrate (1.6 mm thickness) which has a dielectric constant of 4.4. More than 1500 iterations have been carried out during simulation stage to yield the optimized performance characteristics for resonant frequency and return loss at both the above frequencies. Fairly good return losses and radiation pattern response of the proposed dual monopole dual frequency antenna indicates that this antenna is well suited for the applications intended.**

Index Terms—**Microstrip antenna, dual band, impedance bandwidth, monopole antenna, mobile communication.**

I. Introduction

Dual band antenna designs are gaining much interest due to the rapid developments in wireless communication industry. Wire antennas are the oldest and in many case the most versatile antennas from a practical point of view [1]. In modern personal wireless applications, it is desirable to integrate the antenna on a circuit board for low cost, low profile, and conformability. **In the recent years, printed circuit antennas have been receiving much attention owing to their low profile, light weight, low cost, small size, design flexibility, and ease of installation [2]. Applications in the present day mobile communication systems usually require smaller antenna size in order to meet the miniaturization requirements of mobile units. Thus size reduction is becoming major design consideration for practical applications of printed circuit antenna.** In applications in which the increased bandwidth is needed for operating at two separate sub-bands, a valid alternative to the broadening of total bandwidth is represented by dual-frequency patch antennas. **Various types of antennas are available to meet these requirements. Among these, the Dual Monopole Microstrip antenna for Dual Frequency Mobile communication and Wireless LAN Applications has gained great prominence because of simplicity [3]. Dual-frequency antennas exhibit a dual-resonant behavior in a single radiating structure. The antennas designed to operate at 900MHz and 1800MHz are available in literature[4,5,6,7]. This paper discusses the dual frequency antenna operating at 1800MHz and 2400MHz.All the corresponding details of antenna specifications, simulation details and design parameters are clearly presented and discussed.**

II. Antenna Design

The geometry of the microstrip fed uniplanar monopole antenna is shown in Fig. 1, it consists of two printed strip monopoles of different dimensions. The longer one (i.e., Monopole 1) is for a lower frequency band while the shorter one (i.e., Monopole 2) is dominant at the higher frequency. The two monopoles are combined at their lower ends and fed by a single microstrip line with a tuning stub. This structure is printed on a dielectric substrate of dielectric constant 4.4. and thickness 1.6 mm. The design is based on [8].

A. Steps for antenna design

1. Calculate the design parameters and generate the model using HFSS Simulation software.
2. Simulate the antenna by varying one parameter at a time for obtaining desired response.
3. Gather simulated hybrid response data using the return loss plot and radiation pattern with Ansoft HFSS software.
4. Analyze simulated results and verify with manufacturing and fabrication tolerances.
5. Fabricate the proposed design using FR4 with dielectric constant of 4.4.
6. Obtain the measured Return Loss response using vector network analyzer and radiation pattern response in Anachoic chamber.

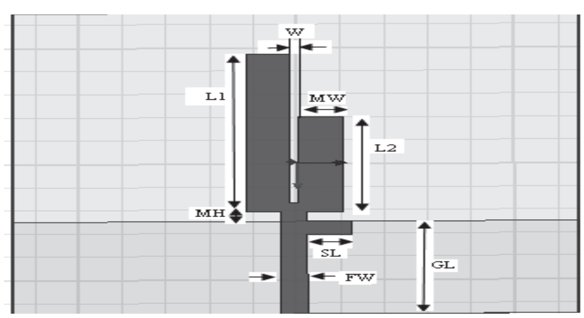


Figure1. Geometry of dual monopole antenna.

7 Compare the simulated and measured results of antenna and analyse the deviations if any.
8 Optimize the antenna performance for aimed specifications.

A microstrip monopole is formed by making one side length of the patch resonant(fraction of wavelength) and other patch dimension very thin. Both the monopoles are chosen to have same width**. The design parameters are calculated using design equations[9].**

$$W = \frac{c}{2f_o\sqrt{\frac{(\varepsilon_r + 1)}{2}}} \quad (1)$$

$$\varepsilon_{reff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{W}\right]^{-\frac{1}{2}} \quad (2)$$

$$\Delta L = 0.412h\frac{(\varepsilon_{reff} + 0.3)\left(\frac{W}{h} + 0.264\right)}{(\varepsilon_{reff} - 0.258)\left(\frac{W}{h} + 0.8\right)} \quad (3)$$

$$L_{eff} = \frac{c}{2f_o\sqrt{\varepsilon_{reff}}} \quad (4)$$

$$L_{eff} = L + 2\Delta L \quad (5)$$

Where W is the width of the patch, ε_r is the dielectric constant of the medium, h is the substrate thickness, f_0 is the resonant frequency of antenna, ε_{reff} is the effective dielectric constant of the medium, h is the height of the substrate, ΔL is the incremental length, L_{eff} is the effective length of the patch. L is the resonant length of the patch.

The width of microstrip line is calculated using the equation,

$$Z_c = \frac{120\pi}{\sqrt{\varepsilon_{eff}}\left[\frac{W_o}{h} + 1.393 + 0.667\ln\left(\frac{W_o}{h} + 1.444\right)\right]} \quad (6)$$

Initially two monopoles are designed separately and the resonance at corresponding frequencies is verified. Both the monopoles are then combined, as for a multiband operation, a multi element monopole is needed [4]. **A dual-band uniplanar monopole antenna operating at 1800 MHz and 2400 MHz has been designed using HFSS Simulation Software. The parametric analysis is carried out with respect to several design parameters of the antenna to the obtain the best design with best radiation characteristics. Optimal design is obtained by choosing the following design parameters. Length of larger monopole as L1=34mm, length of small monopole as L2=20.5mm, feed width FW=3mm, monopole width MW=5mm which is same for both monopoles, the gap between monopoles is W=1mm, SL= 5mm is used for good impedance matching at both the frequencies, ground plane length chosen as 20mm.**

III. RESULTS

The return loss plot determines the resonant frequency and impedance bandwidth of the antenna. It is seen from the return loss plot depicted in Figure 2.that the antenna resonates effectively at both the frequencies and has good impedance bandwidth. The minimum return loss of -19.8dB and -15.8dB are obtained at 1798MHz and 2409MHz respectively. The -10dB return loss bandwidth at lower resonant frequency is 401MHz (1553MHz-1952MHz) and at higher resonant frequency is 404MHz (2421MHz-2825MHz). The impedance characteristics at two resonances is clearly depicted in Fig.3. The radiation patterns of the antenna in E-plane and H-plane at 1800MHz is shown in Fig. 4. It is seen that E plane pattern is bidirectional and H plane pattern is non directional. Fig.5 shows the principal plane patterns at 2450MHz which is also omnidirectional.

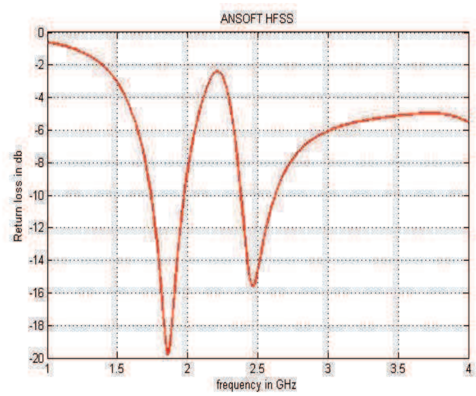


Figure2. Simulated return loss for dual monopole antenna.

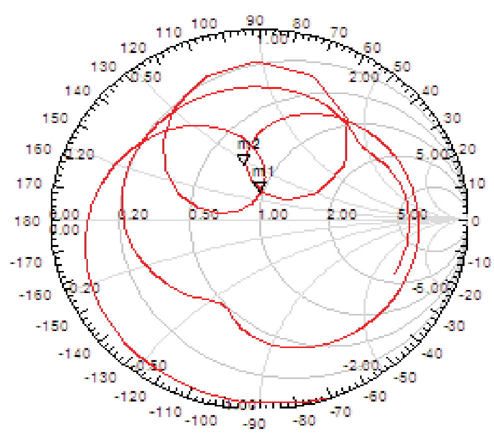


Figure3. Impedance characteristics of the antenna.

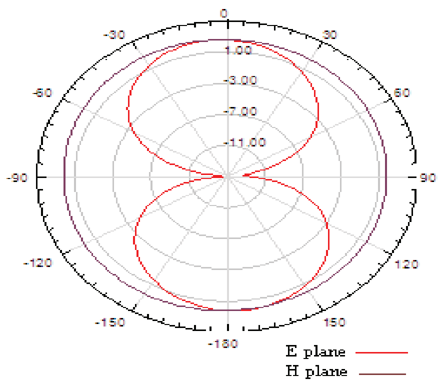


Figure4. Simulated radiation patterns for 1800MHz.

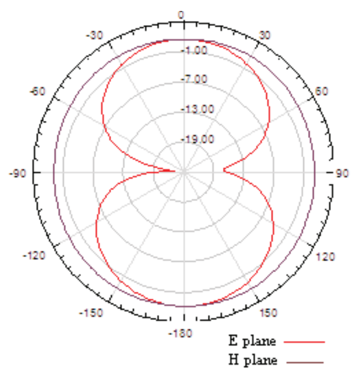


Figure5. Simulated radiation patterns for 2450MHz.

Conclusions

The monopole antenna is very compact and is easily fed at center of two monopoles by a 50 ohm microstrip line. The simulation results of radiation patterns and return loss are presented. The antenna operating characteristics are very good at both the frequencies and are suitable for wireless communication. This simple structure has performance characteristics such as pure vertical polarization and horizontal omnidirectional coverage which makes its extensive use possible in variety of wireless applications.

References

[1] Multiband Integrated Antennas for 4G Terminals by David A.Sánchez-Hernández Universidad Politécnica de Cartagena, Artech House.

[2] Broadband microstrip antennas by Girish kumar and K .P Ray, Artech House.

[3] Microstrip antennas by I.J. Bahl and P.Bhartia , Artech House.

[4] Chen, H.-M., et al., “A Compact Dual-Band Microstrip-Fed Monopole Antenna,” Proc. IEEE Antennas and Propagation Society Int. Symposium, July 2001, Vol. 2, pp. 124–127.

[5] B. Winter and V. Stoiljkovic, "A novel dual antenna for mobile communications," in 1998.

[6] I.Egorov and Z. Ying, "A non-uniform helical antenna for dual-band cellular phones," in 2000 IEEE Antennas Propagat. Soc. Int. Symp. Dig., Salt Lake City, UT, July 2000, pp. 652-655.

[7] S. Tarvas and A. Isohatala, "An internal dual-band mobile phone antenna," in 2000 IEEE Antennas Propagat. Soc. Int. Symp. Dig., Salt Lake City, UT, July 2000, pp. 266-269.

[8] M. Ali, M. 0. Wski, M. A. Stuchly and S. S. Stuchly, "A dual-frequency strip-sleeve monopole antenna for a laptop computer," in 1998 IEEE Antennas Propagat. Soc. Int. Symp. Dig., Atlanta, GA, June 1998, pp. 794-797.

[9] Balanis, C. A., Antenna Theory: Analysis and Design, New York: John Wiley & Sons, 2005.

[10] Nakano, H., et al., “An Inverted FL Antenna for Dual-Frequency Operation,” IEEE Transactions on Antennas and Propagation, Vol. 53, No. 8, August 2005, pp. 2417– 2421.

Low Power Design for CMOS Circuits

P. Kalyani
CVR College of Engineering, Department of ECE, Ibrahimpatan, R.R.District, A.P., India
Email: kalyani.pala06@gmail.com

***Abstract*— Designing high-speed low-power circuits with CMOS technology has been a major research problem for many years. The increasing demand for low-power design can be addressed at different design levels, such as software, architectural, algorithmic, circuit, and process technology level . This paper presents different approaches to reduce power consumption of any arbitrary combinational logic circuit by applying power minimization techniques at circuit level.**

***Index Terms*—Static power, Dynamic power, Dual Vt, Multi threshold voltage,Stacking,Forced stack, Low power design.**

I. Introduction

In the past, the major concerns of the VLSI circuit designers were area, speed and cost. In recent years, this has changed dramatically and power dissipation is being given increased weightage in comparison to area and speed design metrics. Power wall is a clear and present roadblock in the semiconductor industry. The proliferation of portable and hand-held electronics combined with increasing packaging costs is forcing circuit designers to adopt low power design methodologies. Low power design of application specific integrated circuits (ASIC) result in increased battery life and improved reliability. Indeed, the Semiconductor Industry Association technology roadmap has identified low power design techniques as a critical technological need. Hence it becomes imperative for circuit designers to acknowledge the importance of limiting power consumption and improving energy efficiency at all levels of the design hierarchy, starting from the lower levels of abstraction, when the opportunity to save power is significant.

The remaining part of this paper is organized as follows. Section 2 highlights the reasons underpinning the choice of the logic style considered and its typical advantages. Section 3 deals with the issue of power dissipation in CMOS circuits. Section 4 surveys on power minimization techniques and finally conclude.

II. Complementary CMOS Logic Style

Complementary CMOS logic or Static CMOS logic style consisting of complementary nMOS pull-down and pMOS pull-up networks to drive ‘0’ and ‘1’ outputs are used for the vast majority of logic gates in digital integrated circuits. They have good design margins, fast, low power, insensitive to device variations, easy to design, widely supported by commercial CAD tools, and readily available in standard cell libraries. When noise does not exceed the margins, the gate eventually will settle to the correct logic level. Indeed many ASIC methodologies allow only complementary CMOS circuits. Even custom designs use static CMOS for 95% of the logic. They also enable low leakage designs owing to their inherent flexibility to accommodate leakage control transistors at the junction between the pull-up and pull-down network nodes .

Other advantages of static CMOS logic style are its robustness against voltage scaling and transistor sizing and thus ensuring reliable operation at low voltages and arbitrary transistor sizes. Input signals are connected to transistor gates only, which facilitates the usage and characterization of logic cells. The layout of CMOS gates is straightforward and efficient due to the complementary transistor pairs. Given the correct inputs, it will eventually produce the correct output so long as there were no errors in logic design or manufacturing. Static CMOS logic also has the advantage that there is no precharge/predischarge operation and charge sharing does not exist. Other circuit families tend to become prone to numerous pathologies, including charge sharing, leakage, threshold drops and ratioing constraints. Basically, CMOS fulfils all the requirements regarding the ease-of-use of logic gates.

III. Power Dissipation In CMOS Circuits

Average power dissipation (Pavg) in CMOS digital circuits can be expressed as the sum of three main components , which are summarized in the following equation, as

$$P_{avg} = P_{short\text{-}circuit} + P_{leakage} + P_{dynamic}$$

Pshort-circuit is the power from stacked P and N devices in a CMOS logic gate that are in the ON state simultaneously. This happens briefly during switching. This type of power dissipation can be controlled by minimizing the transition times on nets. It usually accounts for 15%-20% of the overall power dissipation.

Pleakage is the power dissipation due to spurious currents in the non-conducting state of the transistor. This component becomes a larger problem as device geometries shrink and transistor threshold voltages (Vt) drop. Leakage current depends upon the supply, Vdd (or how close it is with respect to Vt), Vt itself, transistor aspect ratio (W/L) and temperature. As the supply voltage scales down with technology, this increases exponentially and is construed to dominate the total power dissipation in ultra deep submicron technologies. Increasing die area also increases the leakage power adversely, as this increases the number of transistors.

Pdynamic is the dynamic power dissipation, also called the switching power. This is the dominant source of power consumption in CMOS system-on-chip (SoC), accounting for roughly 75% of the total. It is generally represented by the following approximation,

$$P_{dynamic}=\alpha.C_L.V_{dd}^2.f_{clk} \quad (2)$$

Where 'α' is the switching activity factor (also called transition probability) and it tends to increase as the need for bandwidth increases, 'CL' is the overall capacitance to be charged and discharged in a reference clock cycle. Technology scaling has resulted in smaller transistors and hence smaller transistor capacitances, but interconnect capacitance has not scaled much with process and has become the dominant component of capacitance. 'Vdd' is the supply voltage. Though voltage scaling has the biggest impact on power dissipation (nearly quadratic savings in power), this generally comes at an expense of an increase in delay. 'fclk' is the switching frequency of a global clock for a globally synchronous design, local clock for a locally synchronous design or the input arrival rate in case of a pure static system.

IV. Power Minimization Techniques

Power consumption in a static CMOS circuit basically comprises three components: dynamic switching power, short circuit power and static power. Compared to the other two components, short circuit power normally can be ignored in submicron technology.

A. Dynamic Power

Dynamic power is due to charging and discharging the loading capacitances.

B. Leakage power

The leakage current of a transistor is mainly the result of reverse-biased PN junction leakage, subthreshold leakage and gate leakage as illustrated in Figure 1.

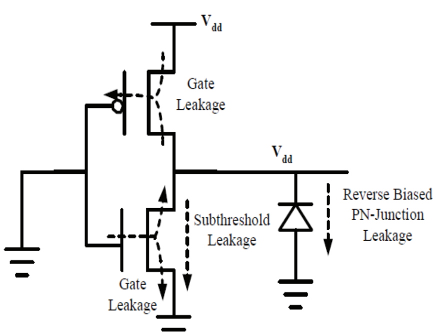


Figure 1 :Leakage currents in an inverter.

Leakage is becoming comparable to dynamic switching power with the continuous scaling down of CMOS technology. To reduce leakage power, many techniques have been proposed, including dual-*Vth*, multi-*Vth*, optimal standby input vector selection, transistor stacking and body bias.

1. Dual-Vth Assignment

Dual-*Vth* assignment is an efficient technique for leakage reduction. In this method, each cell in the standard cell library has two versions, low *Vth* and high *Vth*. Gates with low *Vth* are fast but have high subthreshold leakage, whereas gates with high *Vth* are slower but have much reduced subthreshold leakage. Traditional deterministic approaches for dual-threshold assignment utilize the timing slack of non-critical paths to assign high *Vth* to some or all gates on those non-critical paths to minimize the leakage power.

2. Multi-Threshold-Voltage CMOS

A Multi-Threshold-Voltage CMOS (MTCMOS) circuit is implemented by inserting high *Vth* transistors between the power supply voltage and the original transistors of the circuit. The original transistors are assigned low *Vth* to enhance the performance while high-*Vth* transistors are used as sleep controllers. In active mode, SL is set low and sleep control high-*Vth* transistors (MP and MN) are turned on. Their on-resistance is so small that VSSV and VDDV can be treated as almost being equal to the real power supply. In the standby mode, SL is set high, MN and MP are turned off and the leakage current is low. The large leakage current in the low-*Vth* transistors is suppressed by the small leakage in the high-*Vth* transistors. By utilizing the sleep control high-*Vth* transistors, the requirements for high performance in active mode and low static power consumption in standby mode can both be satisfied.

3. Transistor Stacking

The two serially-connected devices in the off state have significantly lower leakage current than a single off device. This is called the stacking effect. With transistor stacking by replacing one single off transistor with a stack of serially-connected off transistors, leakage can be significantly reduced. The disadvantages of this technique are also obvious. Such a stack of transistors causes either performance degradation or more dynamic power consumption.

4. Optimal Standby Input Vectors

Subthreshold leakage current depends on the vectors applied to the gate inputs because different vectors cause different transistors to be turned off. When a circuit is in the standby mode, one could carefully choose an input vector and let the total leakage in the whole circuit to be minimized model leakage current by means of linearized pseudo-Boolean functions.

Conclusions

When scaling down to deep sub-micron technology leakage is a critical problem. In this paper we have proposed some of the Power minimization techniques.

Experiments have to be conducted on these technique to decide which technique is best suited for power minimization.

References

[1] A. Abdollahi, F. Fallah, and M. Pedram, "Leakage Current Reduction in CMOS VLSI Circuits by Input Vector Control," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 12, no. 2, pp. 140-154, 2004.

[2] M. Borah, R. M. Owens, and M. J. Irwin, "Transistor Sizing for Low Power CMOS Circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 15, no. 6, pp. 665-671, 1996.

[3] A. P. Chandrakasan and R. W. Brodersen, *Low Power Digital CMOS Design*. Boston: Kluwer Academic Publishers, 1995.

[4] F. Gao and J. P. Hayes, "Total Power Reduction in CMOS Circuits via Gate Sizing and Multiple Threshold Voltages," in *Proc. Design Automation Conference*, 2005, pp. 31-36.

[5] M. C. Johnson, D. Somasekhar, C. Lih-Yih, and K. Roy, "Leakage Control With Efficient Use of Transistor Stacks in Single Threshold CMOS," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 10, no. 1, pp. 1-5, 2002.

[6] J. T. Kao and A. P. Chandrakasan, "Dual-Threshold Voltage Techniques for Low-Power Digital Circuits," *IEEE Journal of Solid-State Circuits,* vol. 35, no. 7, pp. 1009-1018, July 2000.

FPGA Implementation of Variable Digital Filter using MicroBlaze Processor

M. Rupa[1] and Rayabandi Karthik Kumar[2]
[1]CVR College of Engineering, Department of ECE, Ibrahimpatan, R.R.District, A.P., India
Email: arupa_ece@yahoo.com
[2] CVR College of Engineering, Department of ECE, Ibrahimpatan, R.R.District, A.P., India
Email: karthikrayabandi@gmail.com

***Abstract*—Digital filters have found numerous applications in most digital signal processing systems. Conventional digital filter can only obtain fixed frequency domain characteristics at a time. In order to obtain variable characteristics, the digital filter's type, number of taps and coefficients should be changed constantly such that the desired frequency-domain characteristics can be obtained. Unlike conventional digital filter, Variable Digital Filters (VDFs) can change their filter-type, number of taps and coefficients constantly such that the desired frequency-domain characteristics can be obtained. This paper proposes a method for Variable Digital Filter (VDF) design based on Field Programmable Gate Array and Embedded Micro-Processor (EMP). This VDF is best suited for realization of digital filter algorithms, which are low-pass, high-pass, band-pass and band-stop filter algorithms with variable frequency domain characteristics. The design has been prototyped on an XUPV5-LX110T Evaluation Platform using Integrated Synthesis Environment (ISE) 12.4 Tools all in one design suit from Xilinx.**

***Index Terms*—Variable Digital Filter; Field Programmable Gate Array; Embedded Micro-Processor**

I. Introduction

With the recent rapid advances in communication systems and real-time signal processing, there has been a constant interest in the design and implementation of digital filters with variable frequencydomain characteristics such as variable cutoff frequency, adjustable passband width. The digital filters with variable frequency domain characteristics are referred to as Variable Digital Filters.

Digital filters are mainly two types- Infinite Impulse Response (IIR) and Finite Impulse Response (FIR) digital filter. The response of an IIR filter is a function of current and past input signal samples and past output signal samples. The dependency on past outputs (i.e., recursive) gives rise to the infinite duration of the filters output response even when the input values have been stopped. A Finite Impulse Response (FIR) filter is a filter structure that can be used to implement almost any sort of frequency response digitally. An FIR filter is usually implemented by using a series of delays, multipliers, and adders to create the filter's output. If the output samples of the system depend only on the present input, and a finite number of past input samples, then the filter has a finite impulse response.

In the most general case of a variable filter, all filter magnitude parameters might be subject to changes. Variable magnitude deviations δ_s (in the stopband) and δ_p (in the passband) are, however, not easily achievable – complicated recalculations of the entire transfer function are required and usually many circuit elements must be changed in order to obtain the new δ_s and δ_p.

Fortunately, more often, only frequency parameters – cutoff frequency w_c and stopband edge w_s of low-pass (LP) and high-pass (HP) filters or w_{c1}, w_{c2}, w_{s1}, w_{s2}, center frequency w_0 and bandwidth BW of band-pass (BP) and band-stop (BS) filters – are tuned.

The multipurpose FIR filter is designed using windowing technique. The windowing method is the preferred algorithm because it is simplest method of FIR filter implementation, the selection of the window is based on the causality and stability of all the filter types, low-pass, high-pass, band-pass, and band-stop.

Most of the multipurpose FIR filter designs use commercial tools such as MATLAB's Filter Design and Analysis (FDA) tool for its implementation. MATLAB's fir1 and fir2 function is utilized to compute the filter coefficients [1]. The implementation of such multipurpose FIR filter starts with calculation of coefficients using MATLAB, followed by HDL program and its implementation [1] [2].

This paper proposes a method for implementation of variable digital filter, eliminating the usage of MATLAB functions for coefficient calculations, by using embedded soft RISC processor core. The processor core is optimized for implementation in XILINX FPGAs, and supports both on-chip Block RAM and external memory.

II. Basic Principle Of FIR Filter

The transfer function for an FIR filter of length N is given as:

$$H(z)=\sum_{k=0}^{N-1} h(k)z^{-k}$$

The design of a digital filter involves five steps[3]:

1. Filter Specification:

This may include stating the type of the filter, for example lowpass filter, the desired amplitude and/or phase responses and the tolerances prepared to accept, the sampling frequency, and the wordlength of the input data.

2. Coefficients calculation:

At this step, one determines the coefficients of a transfer function, H(z), which will satisfy the specifications given in step 1. The choice of the coefficient calculation method will be influenced by several factors, the most important of which are the critical requirements of step 1.

3. Representation of the filter by a suitable structure (realization):

This involves converting the transfer function obtained in step 2 into suitable filter network or structure.

4. Analysis of the effects of finite word length on filter performance:

Here the effects of quantizing the filter coefficients and the input data as well as the effect of carrying out the filtering operation using fixed wordlength on the filter performance are analyzed.

5. Implementation of filter in software and/or hardware:

This involves producing the software code and/or hardware and performing the actual filtering.

III. Implementation Of Variable Digital Filter

The implementation of Variable Digital Filter consists of

- MicroBlaze (Embedded Micro-Processor)
- User-machine interface
- Shared memory
- Programmable FIR Filter

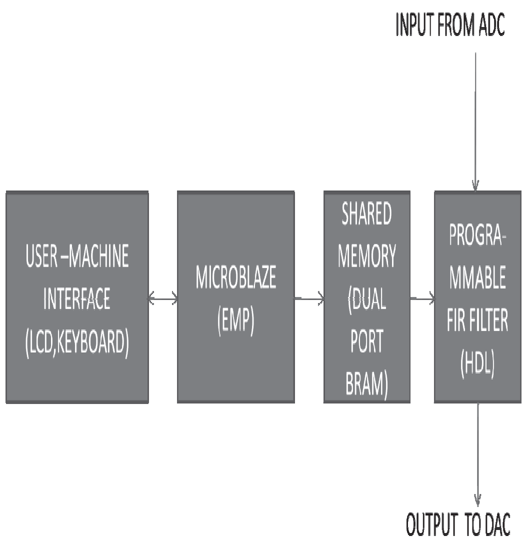


Figure 1. Block Diagram of VDF

A. MicroBlaze (Embedded Micro-Processor)

The *MicroBlaze* soft core processor is used for the computation of coefficients of Variable Digital Filter as per the user specifications. The Embedded Development Kit [4] [5] is used to create a hardware design composed of IP cores and a MicroBlaze soft processor. The design is completed by writing a software application to run on the MicroBlaze processor. The software application controls the functionality of different IP cores added to the processor. The C programming language is used for developing the software application. The below figure 2 shows the block diagram of a MicroBlaze processor based system. The IP cores are connected with MicroBlaze processor using Processor Local Bus v4.6.

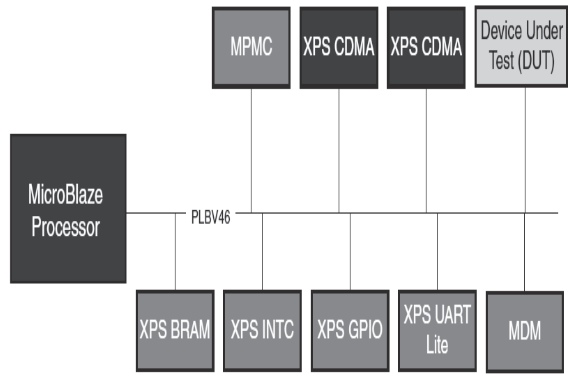


Figure 2. MicroBlaze Processor System[4]

The MicroBlaze processor acts as one of the sub-module in the implementation of VDF.

B. User-machine interface

The implementation of variable Digital Filter uses Liquid crystal display (LCD) and PS/2 keyboard as user-machine interface modules. The LCD module is used to guide the user, such that the required filter specifications can be provided to the processor by using PS/2 keyboard module. The LCD and PS/2 keyboard modules are interfaced with MicroBlaze processor. The MicroBlaze processor monitors the functionality of these modules.

C. Shared memory

A ***Dual Port Block Random Access Memory*** (DP-BRAM) is used to store the coefficient values computed by the MicroBlaze processor. Dual Port BRAM is implemented by adding Block RAM (BRAM) IP Block and XPS BRAM Controller IP to the MicroBlaze system.

The BRAM Block is a configurable memory module that attaches to a variety of BRAM Interface Controllers. Both Port A and Port B of the memory block can be connected to independent BRAM Interface Controllers: Local Memory Bus, Processor Local Bus, and On-Chip Memory.

The XPS BRAM Interface Controller is a Xilinx IP module that incorporates a PLB V4.6 interface. This controller is designed to be byte accessible. Any access size (in bytes) up to the parameterized data width of the BRAM is permitted. The XPS BRAM Interface Controller is the interface between the PLBV4.6 and the BRAM block peripheral. A BRAM memory subsystem consists of the controller along with the actual BRAM components that are included in the BRAM block peripheral.

The filter coefficients computed by the MicroBlaze processor are in the IEEE-754 single precision format, which are stored in the memory using Port A of DP-BRAM.

D. IEEE-754 Single Precision Format

MicroBlaze processor supports single precision IEEE 754 format for representing floating point values. IEEE 754 format is most common standard for representing floating point numbers.

Single precision: 32 bits, consisting of

1. Sign bit (1 bit): The sign bit is as simple as it gets. 0 denotes a positive number; 1 denotes a negative number. Flipping the value of this bit flips the sign of the number.

2. Exponent (8 bits): The exponent field needs to represent both positive and negative exponents. To do this, a *bias* is added to the actual exponent in order to get the stored exponent. For IEEE single-precision floats, this value is 127. Thus, an exponent of zero means that 127 is stored in the exponent field.

3. Mantissa (23 bits): The *mantissa*, also known as the *significand*, represents the precision bits of the number. It is composed of an implicit leading bit and the fraction bits.

The real value assumed by a given 32 bit **binary32** data with a given biased exponent **e** and a **23 bit fraction** is = $-1^{\text{sign}} (1.\, b_{-1}\, b_{-2} \ldots b_{-23})_2 \text{ X } 2^{(e-127)}$.

E. Programmable-FIR Filter

A 15-tap Programmable-FIR filter is developed as a sub-module using Verilog HDL language in Xilinx ISE. The FIR filter is named as programmable; because it can change its filter order, type as per the data received from the MicroBlaze processor. Optimized form of FIR filter structure is used for developing the P-FIR filter. The Port B of DP-BRAM is used by the P-FIR filter to read the data from memory. The order of the P-FIR filter can be changed in odd values. A Push button is used to initialize the P-FIR, after the execution of MicroBlaze sub-module.

F. Process Description

An 8 bit data from the output of an ADC is taken as input of the P-FIR filter. The filter coefficients read from the memory are in IEEE 754 single precision format. To simplify the filter computation process, the IEEE 754 single precision value is converted into a Fixed-Point value (Q11.12) using Floating Point IP. Q11.12 represents 1 sign bit, 11 bit integer value and 12 bit fractional value. The multiplications and addition operations are carried according to the optimized structure of the filter. The 12 Least Significant Bits (LSB) are truncated from the final result. The filter output can be used for further signal processing.

A 125 MHz clock is provided as clocking resource to the MicroBlaze processor by the clock generator core using 100 MHz user clock. The Programmable FIR filter is sourced with 33 MHz on-board clock.

IV. RESULTS

The results obtained by implementing Variable Digital Filter using MicroBlaze on Virtex-5 board are shown. The MicroBlaze processor sub-module is executed at the beginning. This sub-module takes filter specifications entered by the user as inputs and computes the respective filter coefficients.

Table I. compares the MicroBlaze computed filter coefficient values with cutoff frequency = 1.5 KHz, sampling frequency = 8 KHz and filter length = 7 for LPF.

TABLE I.
Comparison of filter coefficient values for LPF

Filter coefficients	Low-Pass Filter			
	Theoretical		Obtained	
	Floating-point	Single-precision	Single-precision	Floating-point
h[0]	-0.077381	BD9E79EE	BD9E79E9	-0.07738096
h[1]	0.212798	3E59E7B8	3E59E7A0	0.21279764
h[2]	0.793155	3F4B0C35	3F4B0C31	0.7931548
h[3]	1.083333	3F8AAAA8	3F8AAAA9	1.0833331

The Programmable-FIR filter sub-module is initiated after the execution of MicroBlaze sub-module. This sub-module reads coefficient values computed by the MicroBlaze sub-module from DP-BRAM to configure the filter structure. An 8-bit data is provided as input to the P-FIR filter module. The 24 bit output of P-FIR filter is used for further signal processing. Figure 3 shows the simulation results of P-FIR filter.

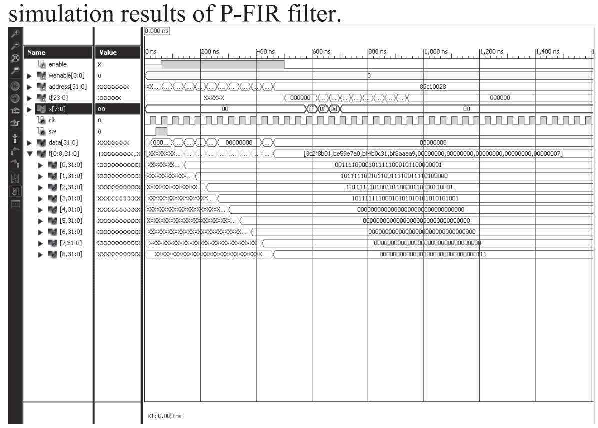


Figure 3. Simulation Result of HPF Filter for fc=1.5 KHz, fs=8 KHz & N = 7

CONCLUSION

Today many embedded products place their solution on several chips making it bigger, more expensive and more power requiring. FPGA boards has become bigger, faster and cheaper and is now able to handle a SoC solution. As the FPGA boards have become bigger and faster they are now able to handle a soft processor which is an Intellectual Property (IP) core implemented using logical primitives. A key benefit is configurability where it is possible to add only what is needed in the design. A trade off is performance, a hard processor is faster but less configurable and more expensive.

With a soft-core processor, a design becomes more flexible, while still keeping high performance parts inside the FPGA. Thus the Xilinx MicroBlaze 32 bit RISC soft-core processor can be easily adapted to the designer's needs and full- featured GNU tool chain is used for software development.

Variable Digital Filter using MicroBlaze processor has been implemented on Virtex-5LX110T board. The processor is operated at a clock frequency of 125 MHz. Variable Digital Filter is assigned to a frequency of 33MHz, and the maximum speed of operation that can be achieved is 56MHz.

REFERENCES

[1] Md.Syedul Amin, Md.Mamun, Labonnah F. Rahaman,"Hardware Approach of a Multipurpose Finite Impulse Response Filter for Real-Time Filtering Applications", American Journal of Applied Sciences.

[2] R.Yousefi, A.Ahmadi, S.M.Fakhraie, "Design Flow for Hardware Implementation of DigitalFilters", International Symposium on Telecommunications, 2008.

[3] [3] E.C.Ifeachor, B.W.Jervis, "Digital Signal Processing-A Practical approach", Prentice Hall, 2nd Edition, 2002.

[4] MicroBlaze Processor Reference Guide *Embedded Development Kit EDK 12.4*

[5] Embedded System Tools Reference Manual EDK 12.4

Reverse Ad Hoc on Demand Distance Vector Routing Protocol in MANETs and Performance Comparison with AODV

Shakeel Ahmed
CVR College of Engineering, Department of ECE, Ibrahimpatan, R.R.Distrcit, A.P., India
Email: shakeel_be@yahoo.com

***Abstract*— An ad hoc wireless network consists of a set of mobile nodes that are connected by wireless links. The network topology in such a network may keep changing randomly. Routing protocols that find a path to be followed by data packets from source node to a destination node used in traditional wired networks cannot be directly applied in ad hoc wireless networks due to their highly dynamic topology. A variety of routing protocols for ad hoc wireless networks has been proposed in the recent past. AODV (Ad-hoc on-demand Distance vector routing) is a representative among the most widely studied on-demand ad hoc routing protocols. In AODV, source node floods the route request packet in the network to obtain a route for the destination. AODV and most of the on demand ad hoc routing protocols use single route reply along reverse path. Route reply packet may not reach the source node due to rapid changes in topology or link failure resulting in reinitiating route discovery process. This increases communication delay, power consumption & decreases packet delivery ratio. To avoid these problems, the proposed routing protocol generates & maintains multiple route replies. The proposed (modified) AODV protocol is called Reverse AODV (R-AODV). In this paper, R-AODV protocol is implemented using NS-2.35 in Linux platform. Simulations are conducted to evaluate the performance of R-AODV and is compared with AODV using application oriented metrics, such as the throughput, packet delivery ratio and end to end delay. Simulation results show that R-AODV performs well when link breakage is frequent.**

***Index Terms*—AODV, MANETs, NS-2.**

I. Introduction

Mobile ad hoc network [1] is a dynamic network which allows communication between the mobile nodes without a central administrator. The network topology in such a network may keep changing randomly. A variety of routing protocols [2] for ad hoc wireless networks have been proposed in the recent past. Ad hoc wireless network routing protocols [2] can be classified into three major categories based on the routing information update mechanism.

1. *Proactive or table driven routing protocols*:

In this, each node maintains the network topology information in the form of routing tables by periodically exchanging routing information. Routing information is generally flooded in the whole network. Whenever a node needs a route to the destination it runs an appropriate path finding algorithm on the topology information it maintains.

2. *Reactive or on demand routing protocols*:

Such protocols do not maintain the network topology information. They obtain the necessary route when it is required, by using a connection establishment process. Hence these protocols do not exchange routing information periodically.

3. *Hybrid routing protocols*:

These protocols combine the best features of the above two categories. Nodes with a certain distance from the source node concerned or within a particular geographical region are said to be within the routing zone of the given node. For routing within this zone, a table-driven approach is used. For nodes located beyond this zone, an on-demand approach is used.

Focus of study is on-demand routing protocols. One of the on-demand routing protocol is AODV [3]. The main advantage of this protocol is that routes are established on demand i.e., only when it is required by a source node for transmitting data packets. But due to the dynamic change of network topology, links between nodes are not permanent. When a link breaks, a node cannot send packets to the intended next hop node resulting in packet loss. If the lost packet is a route reply packet it brings much more problems as the source node needs to reinitiate route discovery procedure.

The route discovery procedure and design of AODV protocol is discussed by C. Pekin "et al." in [3]. The design of extended AODV(R-AODV) also called Reverse AODV and the comparative analysis of AODV with R-AODV using UDP traffic for constant bit rate applications considering scalability is discussed by E.Talipov "et al." in [4].

In this paper the performance comparison of the modified AODV (R-AODV) [4] algorithm in which route reply message is multicast to its neighbors resulting in redundant route reply messages instead of unicasting the route reply to its next hop as in the traditional AODV is done. With this the probability of a successful route discovery is increased as we have repetitious route reply messages in our network.

The robustness of the R-AODV algorithm is tested and compared with the existing AODV algorithm by using UDP as traffic source.

The rest of the paper is organized as follows: Section II gives a brief introduction of AODV routing protocol and an overview of modified AODV(R-AODV) routing protocol. Simulation setup is described in section III. Section IV gives the results and performance comparison of the two routing protocols. Section V concludes the paper.

II. Classification Of routing Protocols

a) Ad hoc on demand distance vector (AODV)

Ad hoc on demand distance vector (AODV) [3] routing protocol creates routes on-demand. In AODV, a route is created only when requested by a network connection and information regarding this route is stored only in the routing tables of those nodes that are present in the path of the route. The procedure of route establishment is as follows. Assume that node X wants to set up a connection with node Y. Node X initiates a path discovery process in an effort to establish a route to node Y by broadcasting a Route Request (RREQ) packet to its immediate neighbors. Each RREQ packet is identified through a combination of the transmitting node's IP address and a broadcast ID. The latter is used to identify different RREQ broadcasts by the same node and is incremented for each RREQ broadcast. Furthermore, each RREQ packet carries a sequence number which allows intermediate nodes to reply to route requests only with up-to date route information. Upon reception of an RREQ packet by a node, the information is forwarded to the immediate neighbors of the node and the procedure continues until the RREQ is received either by node Y or by a node that has recently established a route to node Y. If subsequent copies of the same RREQ are received by a node, these are discarded. When a node forwards a RREQ packet to its neighbors, it records in its routing table the address of the neighbor node where the first copy of the RREQ was received. This helps the nodes to establish a reverse path, which will be used to carry the response to the RREQ. AODV supports only the use of symmetric links. A timer starts running when the route is not used. If the timer exceeds the value of the 'lifetime', then the route entry is deleted.

Routes may change due to the movement of a node within the path of the route. In such a case, the upstream neighbor of this node generates a 'link failure notification message' which notifies about the deletion of the part of the route and forwards this to its upstream neighbor. The procedure continues until the source node is notified about the deletion of the route part caused by the movement of the node. Upon reception of the 'link failure notification message' the source node can initiate discovery of a route to the destination node.

b) Modified AODV (R-AODV)

Most of on-demand routing protocols, except multipath routing uses single route reply along the first reverse path to establish routing path. In high mobility, pre-decided reverse path can be disconnected and route reply message from destination to source can be missed. In this case, source node needs to retransmit route request message. AODV protocol uses a single route reply message which may be lost in a network with mobile nodes. Transmission control protocols uses acknowledgements to confirm successful data transmission. When TCP is used as a transport layer protocol in MANET which employs AODV at network layer, it deteriorates the performance of the network when mobility is high. The main purpose of study is to increase the possibility of establishing routing path with less RREQ messages than the other protocol has, when topology changes by nodes mobility.

The modified AODV (R-AODV) [4] protocol discovers routes on-demand using a reverse route discovery procedure. During route discovery procedure source node and destination node plays some role from the point of sending control messages. Thus after receiving RREQ message, destination node floods reverse request (R-RREQ), to find source node. When source node receives an R-RREQ message, data packet transmission is started immediately.

III. Simulation Setup

The R-AODV [4] protocol incorporates a route reply similar to route request in AODV [3]. To verify the hypothesis, R-AODV is implemented by changing the source code of AODV in NS2 simulator [5] to enable multiple route reply packets. The simulation setup is described in Table I.

Parameter	Value
Platform	Ubuntu 11.10
NS version	NS- 2.35
No. of Nodes	10
Simulation Terrain Size	500 m x 400 m
Simulation Time	80 seconds
Application Layer	CBR
Traffic Source	UDP

Table I. Simulation parameters

Validation module is build by constructing a scenario of 10 mobile nodes using TCL script. The awk script is run on the trace file obtained after the simulation in Linux Kernel to obtain the statistics of throughput, delay and packet delivery ratio. Comparison between AODV and R-AODV is made under UDP considering the extracted statistics.

a). Performance Metrics:

- Throughput: Throughput is the total of all packets successfully delivered to destination over total-time and result is found as Kbps.
- Average End-to-End Delay: Indicates how long it took for a packet to travel from the source to the application layer of the destination. Calculated in ms.
- Packet Delivery Ratio (PDR)/Packet Delivery Fraction (PDF): Is the ratio of the number of packets successfully received by all destinations to the total number of packets injected into the network by all sources. This is calculated in terms of percentage.

IV. Simulation Results

In this section a comparative analysis of the performance metrics of both the on demand routing protocols AODV and R-AODV with UDP as the traffic source for 10 mobile nodes is done. Simulations are performed for five times and the average value is calculated in order to achieve accurate results as in mobile ad hoc networks, mobile nodes keep on moving at various times.

Figure 1. shows NAM window with data transfer among two source-destination pairs 0-9 & 1-8.

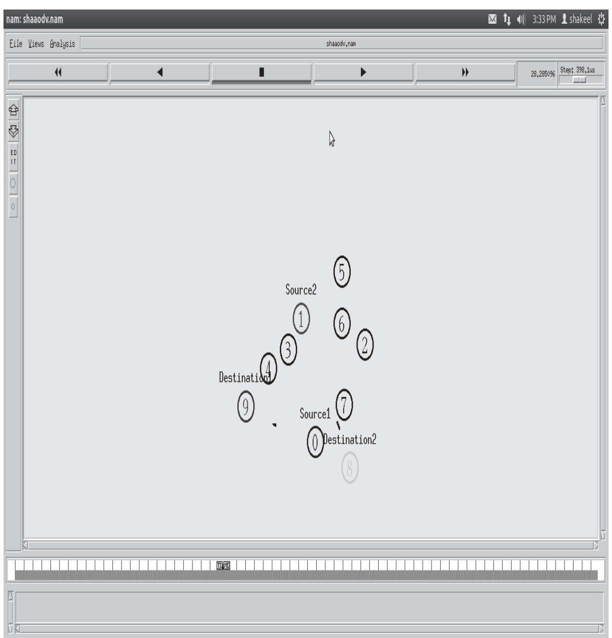
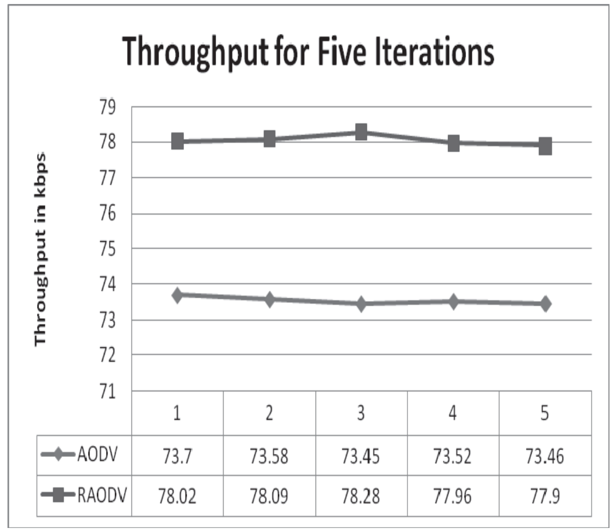


Figure 1.Screenshot of data transfer

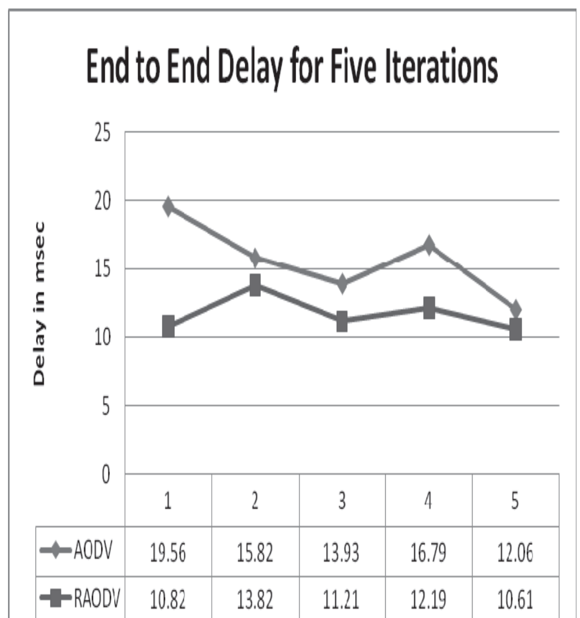
Figure 2. shows throughput obtained for AODV and R-AODV routing protocols for a 10 node network. From figure it is clear that R-AODV gives better throughput performance compared to AODV protocol.



	1	2	3	4	5
AODV	73.7	73.58	73.45	73.52	73.46
RAODV	78.02	78.09	78.28	77.96	77.9

Figure 2. Throughput comparison graph of AODV & R-AODV

End to end delay is more for AODV compared to R-AODV as shown in figure 3. This is due to the fact that route discovery in AODV consumes more time compared to R-AODV and may be reinitiated more number of times.



	1	2	3	4	5
AODV	19.56	15.82	13.93	16.79	12.06
RAODV	10.82	13.82	11.21	12.19	10.61

Figure 3. End to end delay comparison graph of AODV & R-AODV

Figure 4. shows the packet delivery ratio obtained for AODV and R-AODV. From figure it is clear that when an average value is considered then packet delivery ratio is approximately same for both routing protocols.

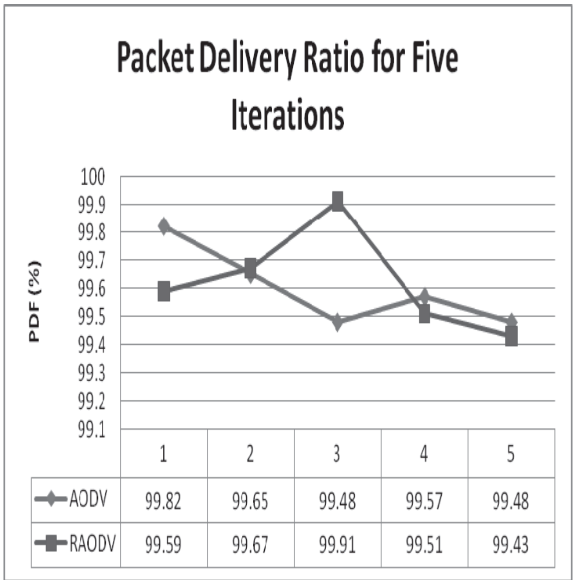


Figure 4. Packet delivery ratio comparison graph of AODV & R-AODV

Table II. Shows the average values of throughput, end to end delay and packet delivery ratio, calculated after five iterations.

No of Executions	Throughput in Kbps		End to End delay in msec		Packet delivery ratio in %	
	AODV	RAODV	AODV	RAODV	AODV	RAODV
1	73.70	78.02	19.56	10.82	99.82	99.59
2	73.58	78.09	15.82	13.82	99.65	99.67
3	73.45	78.28	13.93	11.21	99.48	99.91
4	73.52	77.96	16.79	12.19	99.57	99.51
5	73.46	77.90	12.06	10.61	99.48	99.43
AVG	73.54	**78.05**	15.63	**11.73**	99.60	**99.62**

Table II. Average values

CONCLUSION

Successful delivery of route reply message is very important in a MANET as a lot of route discovery effort is wasted if a reply message is lost, moreover a new route discovery process has to be reinitiated. In the proposed R-AODV protocol frequent route discovery is avoided due to multiple route reply messages resulting in less routing overhead. Thus, from simulation results R-AODV protocol has better throughput and average end to end delay. Further multiple route reply messages in MANET results in approximately same packet delivery ratio.

In this paper the two on-demand routing protocols AODV & R-AODV are analyzed and their performances have been evaluated with respect to three performance metrics using UDP as the traffic source. This paper can be enhanced by analyzing other MANET routing protocols with different traffic sources.

REFERENCES

[1] M.S. Carson, S. Batsell and J. Macker, “Architecture consideration for Mobile Mesh Networking,” Proceedings of the IEEE Military Communications Conference(MILCOM), vol.1, pp 225-229, 21-24

[2] S.Corson and J.Macker, “Routing Protocol performance Issues and Evaluation Considerations”, RFC2501, IETF Network Working Group, January 1999.

[3] C.Perkin, Elizabeth M. Royer, “Ad hoc on demand Distance Vector Routing”, RFC 3561, July 2003, http.//www.ietf.org/rfc/rfc3561.txt

[4] C Kim, E.Talipov, and B.Ahn, A Reverse AODV (R-AODV) Routing Protocol in ad hoc Mobile Networks, in the 2006 IFIP International Conference on Embedded and Ubiquitous Computing” (EUC'06), LNCS 4097, Seoul, Korea, August 2006, pp.522-531.

[5] “The Network Simulator- ns-2”, available at http://www.isi.edu/nsnam/ns/ released on Nov 2011.

Design of an Op-Amp Based Low Voltage Low Dropout Regulator Using 180nm CMOS Technology

Silpakesav[1], Dr.K.S.Nayanatara[2] and Dr.B.K.Madhavi[3]
[1]CVR College of Engineering, Department of ECE, Ibrahimpatan, R.R.District, A.P., India
Email: shilpakesav@gmail.com
[2]CVR College of Engineering, Department of ECE, Ibrahimpatan, R.R.District, A.P., India
Email: ksathirajunayanathara@gmail.com
[3]CVR College of Engineering, Department of ECE, Ibrahimpatan, R.R.District, A.P., India
Email: bkmadhavi2009@gmail.com

***Abstract*—This paper describes the designing of a Low Voltage, Low Dropout (LVLD) Regulator with fast self-reacting (FSR) technique in 0.18 µm CMOS process. The LVLD Regulator provides 1.2 V regulated output voltage. The design procedure includes pass transistor, error amplifier, fast self-reacting path and bandgap reference (BGR) circuit. The pass transistor is designed for a desired minimum dropout voltage. The error amplifier has been developed with a gain of 65 dB and UGB of atleast 1 MHz. Curvature compensated CMOS BGR with 1.8 V supply has been developed to produce 0.8 V output voltage with better temperature coefficient.**

***Index terms*—LVLD, pass transistor, error amplifier, FSR, BGR, curvature compensation.**

I. Introduction

Industry is pushing towards complete system-on-chip (SoC) design solutions among which one is power management. The study of power management techniques has increased spectacularly within the last few years corresponding to a vast increase in the use of portable, handheld battery operated devices. The aim of the power management scheme is to decrease the power consumption by prolonging battery life. Today Voltage regulators are found in nearly every electronic device. They provide the DC voltage for all the electronic circuits used in modern day applications. These applications range from High-speed multiprocessors to multi function cell phones. Voltage regulators can be split into two categories- Linear voltage regulators (LVR) and Switch mode power converters (SMPC).

An LVR can be considered as a building block of every power supply used in microelectronics. A dynamically adjusting resistor is sufficient to regulate the voltage. These convert a noisy input voltage into a stable and clean supply voltage by implementing the dynamically changing resistor pass element. Even though the SMPC offers higher efficiency, the switching noise affects the functionality of the circuits driven by it. The output noise available in commercial LVR is lower than SMPC. LVLD is one such LVR. LVLD regulators are used to provide efficient power management in mobile phones.

The proposed scheme gives less dropout voltage [1], desired line regulation, load regulation, PSRR and output voltage. Section II described the architecture of LVLD regulator with FSR technique, Section III deals with Op-amp designing, Section IV discusses the bandgap reference (BGR) biasing circuit, Section V describes the designing of LVLD regulator, Section VI shows the simulation results of LVLD regulator, the conclusion is given at last.

II. LVLD Regulator with FSR Technique

The block diagram of LVLD regulator with FSR technique is as shown in Fig.1. The key features are fast load transient responses, high loop gain, very low quiescent current and small on chip compensation capacitance.

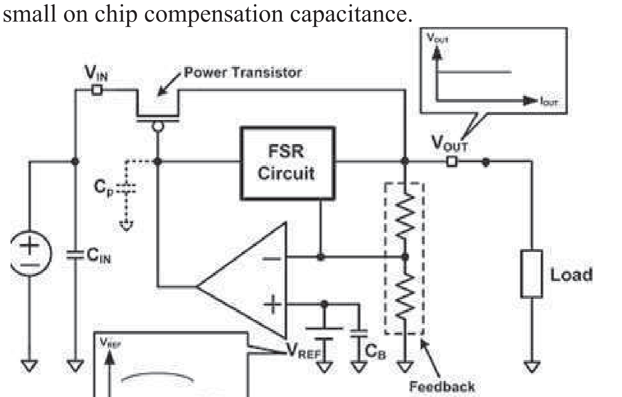


Figure 1. LVLD regulators with fast self-reacting (FSR) technique.

The proposed regulator has one main feedback loop and three ultra fast self-reacting paths. The need for the FSR paths are to decrease the output voltage when load current increases instantly where it achieves the name regulator. The designing of entire LVLD regulator includes pass transistor, feedback resistor, error amplifier and bandgap reference (BGR) circuit.

III. OPERATIONAL AMPLIFIER

The Operational amplifier (op-amp) is the fundamental building block of analog integrated circuit design. Here it is designated to be used as an error amplifier. It is designed to achieve the desired gain and UGB of atleast 1 MHz. The schematic of this op-amp is shown in Fig. 2. The loop gain is achieved from the line and load regulations. Gain of the op-amp is given as,

$$A_v = \frac{2}{(\lambda_n+\lambda_p)^2}\sqrt{\frac{k_n k_p \left(\frac{W}{L}\right)_1\left(\frac{W}{L}\right)_2}{I_{D1} I_{D7}}} \qquad (1)$$

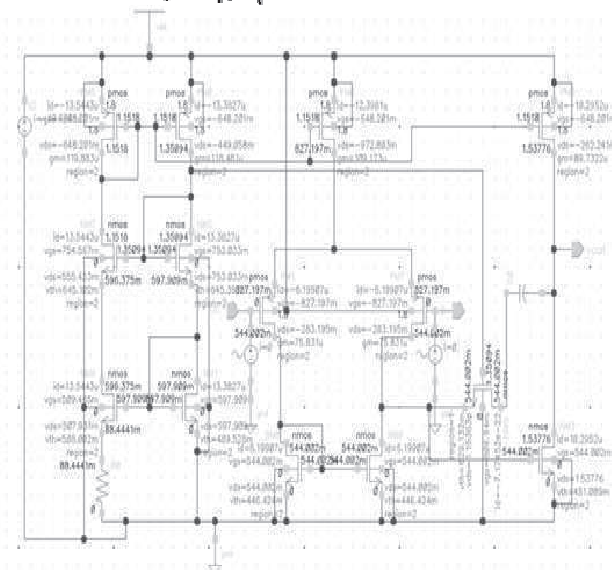


Figure 2. Operational Amplifier

Loop gain is distributed between the error amplifier, pass transistor and feedback resistor ratio. Majority contribution is attributed to the op-amp. The gain of 75 dB and a phase margin of 74.92^o are obtained. The simulated gain plot and phase plot are obtained as shown in Figure 3.

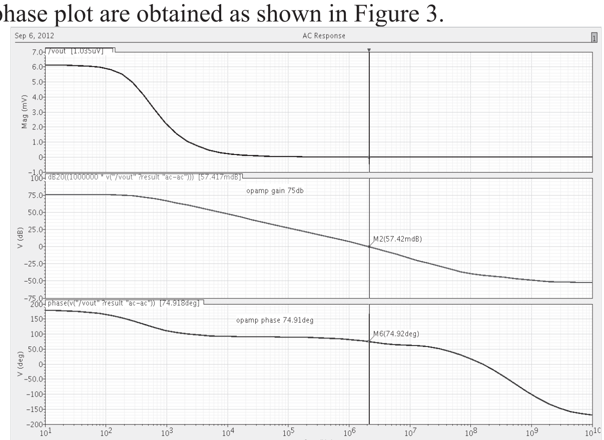


Figure 3. Gain and Phase plot

IV. BANDGAP REFERENCE BIASING

The bandgap voltage reference, which was firstly proposed by Widlar and was further developed by Kuijk and Brokaw, is the one commonly used in many advanced designs and many commercial products since it can provide a predictable reference voltage. The proposed reference scheme can provide from 0 V to near supply voltage by adjusting current mirror and resistor. The schematic of bandgap reference circuit is shown in Figure 4.

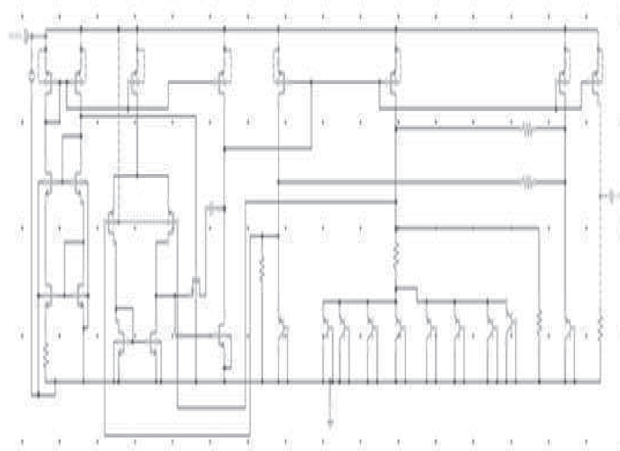


Figure 4. Bandgap Reference circuit

It combines the positive TC[3] of the thermal voltage with the negative TC of the diode forward voltage in a circuit to achieve a voltage reference with a zero TC. The temperature dependence of BGR is shown in Fig. 5.

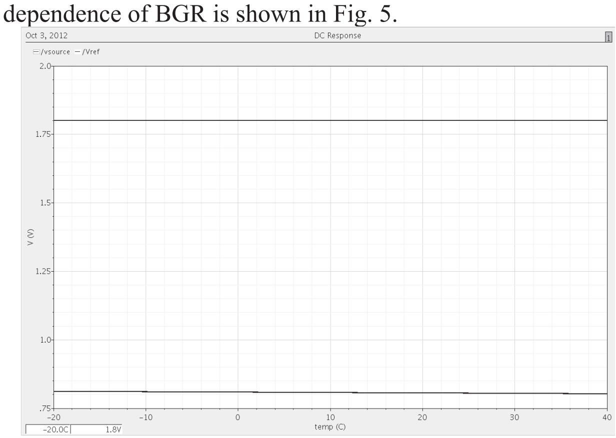


Figure 5. Temperature dependence of BGR

The PSRR plot is shown in Figure 6.

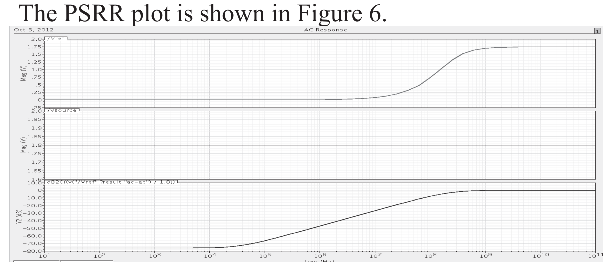


Figure 6. PSRR of BGR

The obtained PSRR is 72 dB at 1 KHz and provides a stable voltage independent of supply, process and temperature.

V. DESIGNING OF LVLD REGULATOR

The pass transistor is designed to deliver a drain current of 100 mA. Careful designing of it is required to have low dropout. The feedback resistor should have a relation one is twice the other. The dropout can be written as

$$V_{dropout} = V_{DSAT} = \sqrt{\frac{2I_{max}}{\mu_p C_{ox}\left(\frac{W}{L}\right)}} \qquad (2)$$

The curvature compensation scheme[3] is used pushes the pole at the gate terminal of the power transistor and pole at the

LVLD regulator output towards higher frequency than the UGF. The overall circuit works like a single pole system and hence has high phase margin with good stability. The schematic of the LVLD regulator is shown in Figure 7.

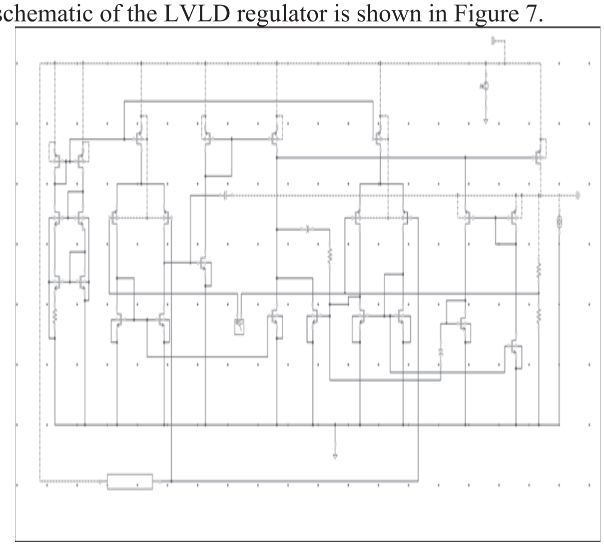


Figure 7: LVLD regulator

VI. SIMULATION RESULTS OF THE LVLD REGULATOR

The frequency response was obtained for zero load current and full load current. The response for zero load current is shown in Fig. 8. The response for full load current of 100 mA is shown in Fig. 9. For full load current LVLD regulator achieves phase margin of 98^o and loop gain of 75 dB. For Zero load current LVLD regulator achieves phase margin of 63^o and loop gain of 65 dB. Fig. 10 shows the line regulation of the LVLD regulator. Fig. 11 shows the load regulation of LVLD regulator. Transient response of LVLD regulator for load current step and line voltage step is shown in Fig. 12 and Fig. 13. The PSRR of LVLD regulator is shown in Fig. 14. Obtained results are shown in Table I.

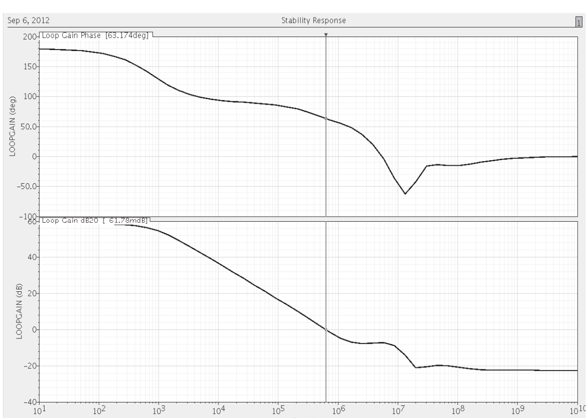


Figure 8: Frequency response for zero load current

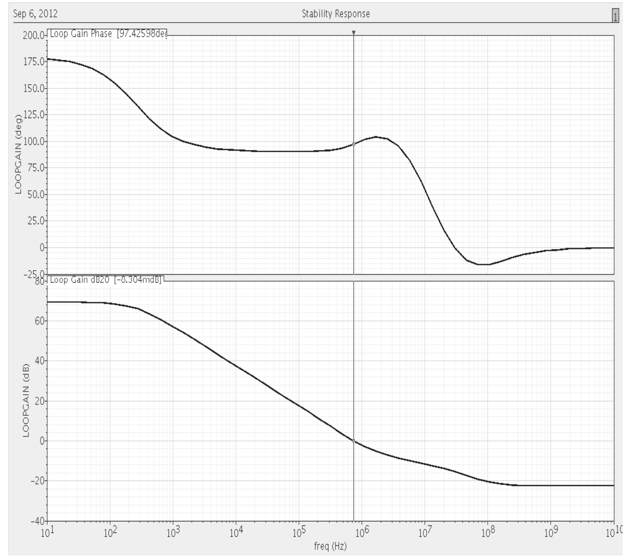


Figure 9: Frequency response for full load current

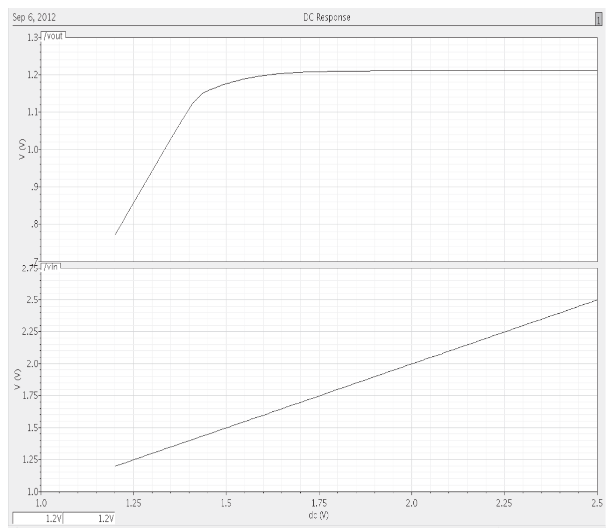


Figure 10: Line regulation of LVLD regulator

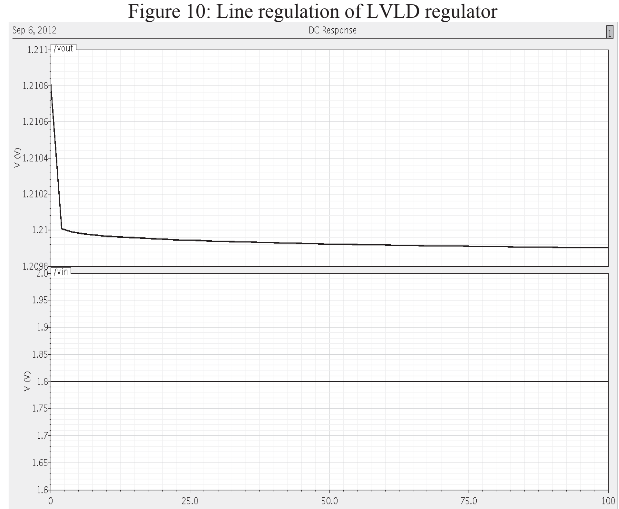


Figure 11: Load regulation of LVLD regulator

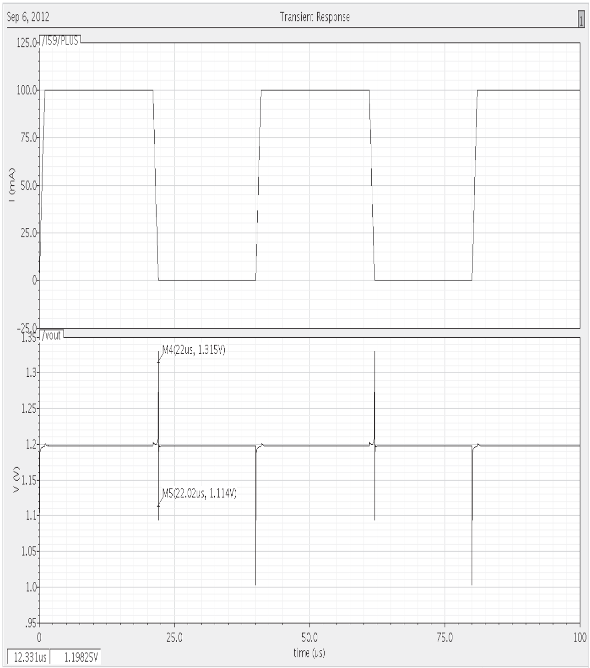


Figure 12: Transient response for load current step

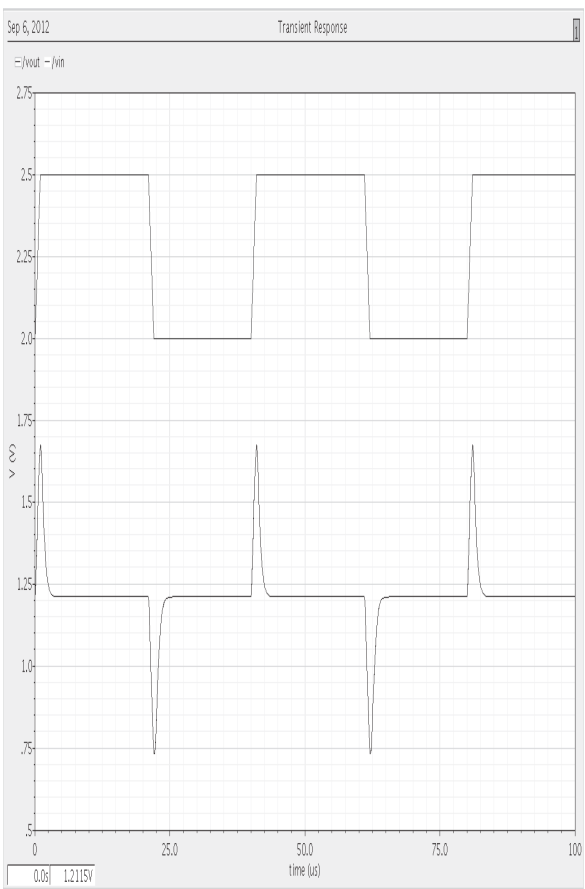


Figure 13: Transient response for line voltage step

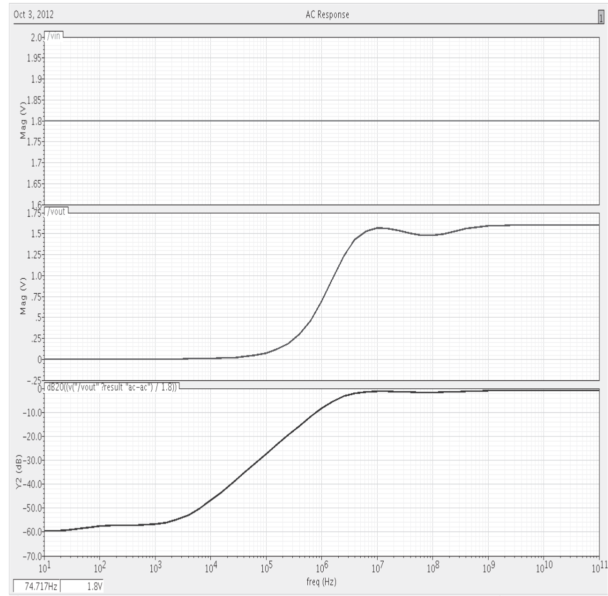


Figure 14: PSRR of LVLD regulator

TABLE I.
OBTAINED RESULTS

Parameter	Obtained value
Technology	**0.18 µm**
$I_{LOAD(max)}$	**100 mA**
Load regulation	**0.162 mV/mA**
Line regulation	**1 mV/V**
PSRR	**52 B**

CONCLUSIONS

In this paper a LVR type Low Voltage Low Dropout (LVLD) Regulator using Fast Self-Reacting (FSR) technique is presented. The proposed reference scheme provides from 0 V to near supply reference voltage. An operational amplifier designed with desired gain and UGB used as an error amplifier. Curvature compensation technique is used for stability of the circuit. The regulator has PSRR about 52 dB, Load regulation of 0.162 mV/mA, Line regulation of 1mV/V.

REFERENCES

[1] G.A. Rincon-Mora and P.E. Allen, “ A low voltage, low quiescent current, low drop-out regulator”, IEEE J. Solid State Circuits, vol.33, pp.36-34, Jan. 1998.

[2] Texas Instruments, “Technical review of low dropout voltage regulator operation and performance”, Texas Instruments Inc., SLVA072, Dallas, TX, USA, August 1999.

[3] “Curvature-Compensated CMOS Bandgap Reference with 1.8-V Operation”Wang Xichuan, Si Cuiying, Xu Xing,Microelectronic Research & Development Center of Shanghai University.

Design and Control of Distributed Power-Flow Controller (DPFC)

R.Naveena Bhargavi
CVR College of Engineering, Department of EEE, Ibrahimpatan, R.R.District, A.P., India
Email: bhargavi.rn5@gmail.com

***Abstract*—This paper represents a new dimension of the Flexible AC Transmission System (FACTS) called as Distributed Power-Flow Controller (DPFC). The aim of this paper is to develop a new type of power-flow controlling device that offers the same control capability as the UPFC, at a reduced cost and with increased reliability. The new device, the so-called Distributed Power Flow Controller (DPFC), is a further development of the UPFC. The DPFC eliminates the common DC link within the UPFC, to enable the independent operation of the shunt and the series converters which enhances the effective placement of the series and shunt converters.**

***Index Terms*—FACTS, Power Flow Controlling Devices, UPFC, DPFC.**

I. Introduction

A. Power Flow Controlling Devices

Power flow is controlled by adjusting the parameters of a system, such as voltage magnitude, line impedance and transmission angle. The device that attempts to vary system parameters to control the power flow can be described as a Power Flow Controlling Device (PFCD)[3].Depending on how devices are connected in systems, PFCDs can be divided into shunt devices, series devices, and combined devices (both in shunt and series with the system), as shown in Figure 1.

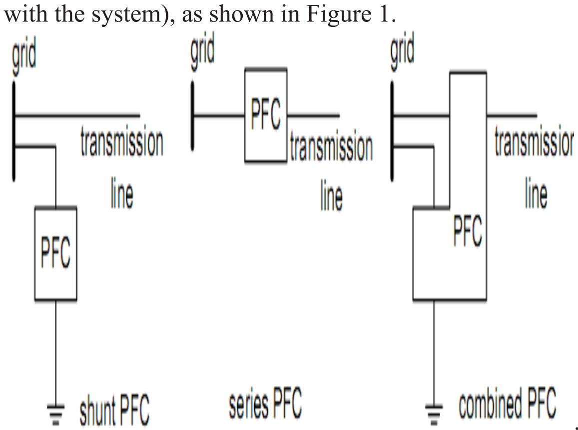


Figure 1. Simplified diagram of shunt, series and combined devices

Based on the implemented technology, PFCDs can be categorized into mechanical-based devices and power electronics (PE)-based devices.

Mechanical PFCDs consist of fixed or mechanical interchangeable passive components, such as inductors or capacitors, together with transformers. Typically, mechanical PFCDs have relatively low cost and high reliability. However, because of their relatively low switching speed (from several seconds to minutes) and step-wise adjustments of mechanical PFCDs, they have relatively poor control capability and are not suitable for complex networks of the future.

PE PFCDs also contain passive components, but include additional PE switches to achieve smaller steps and faster adjustments. There is another term – Flexible AC Transmission System (FACTS) - that overlaps with the PE PFCDs. According to the IEEE, FACTS is defined as an 'alternating current transmission system incorporating power electronic based and other static controllers to enhance controllability and increase power transfer capability'. Normally, the High Voltage DC transmission (HVDC) and PE devices that are applied at the distribution network, such as a Dynamic Voltage Restorer (DVR), are also considered as FACTS controllers. Most of the FACTS controllers can be used for power flow control. However, the HVDC and the DVR are out of the scope of the PFCD. The relationship between the PFCDs, FACTS controllers and mechanical controller is shown in Figure 2.

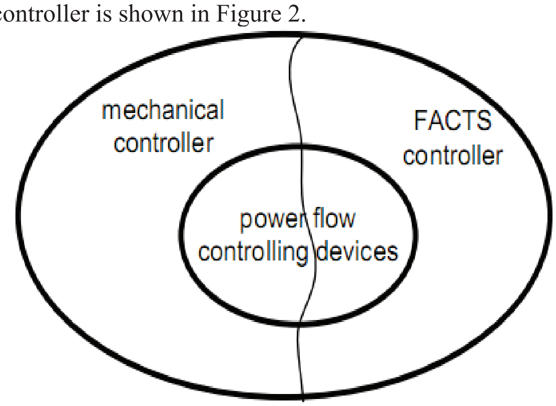


Figure 2. Relationship between the PFCDs, FACTS controllers and mechanical controller

PE PFCD devices can be further subcategorized into two types according to the applied switch technologies: thyristor-based devices and Voltage Source Converter (VSC) based devices. Thyristor PFCDs use inverse-parallel thyristors in series or in parallel with passive components. By controlling the firing angle of the thyristors, the impedance of the device can be adjusted. A thyristor can be controlled to turn on but not to turn off.

It will turn off automatically when the current goes negative. Consequently, the thyristor can only be turned on once within one cycle. The switching frequency of thyristor PFCDs is therefore limited to the system frequency (50/60Hz), resulting in low switching losses. Because thyristors can handle larger voltages and currents than other power semiconductors, the power level of thyristor PFCDs are also higher. The thyristor PFCDs are simpler than VSC PFCDs, allowing them higher reliability. However, the waveforms of voltages and currents generated by thyristor PFCDs contain a large amount of harmonics, thereby requiring large filters.

VSC PFCDs employ advanced switch technologies, such as Insulated Gate Bipolar Transistors (IGBT), Insulated Gate Commutated Thyristors (IGCT), or Metal Oxide Semiconductor Field Effect Transistors (MOSFET) to build converters. Because these switches have turn-on and turn-off capability, the output voltage of a VSC is independent from the current. Consequently, it is possible to turn the switches on and off within the VSC multiple times within one cycle. Several types of VSCs have been developed, such as multi-pulse converters, multi-level converters, square-wave converters, etc. These VSCs proved a free controllable voltage in both magnitude and phase. Due to their relatively high switching frequency, VSC PFCDs make practically instant control (less than one cycle) possible. High switching frequencies also reduce low frequency harmonics of the outputs and even enable PFCDs to compensate disturbances from networks. Therefore, VSC PFCDs are the most suitable devices for future power systems.

On the other hand, there are some challenges facing VSC PFCDs. Firstly, because large amounts of switches are connected in series or in parallel to allow the high voltage and high current through, the VSC PFCDs are expensive. In addition, due to their higher switching frequency and higher on-state voltage in comparison with thyristors, VSC PFCD losses are higher as well. However, with developments in power electronics (such as Silicon-carbide, Gallium-Nitride and synthetic diamond), VSC PFCDs can become more feasible and cost-effective in the future.

According to the above considerations of different types of PFCDs, it can be concluded that PE combined PFCDs (also referred to as combined FACTS) have the best control capability among all PFCDs. They

inherit the advantages of PE PFCDs and combined PFCDs, which is the fast adjustment of multiple system parameters. The Unified Power Flow Controller (UPFC) and Interline Power Flow Controller (IPFC) are currently the most powerful PFCDs; they can adjust all system parameters: line impedance, transmission angle, and bus voltage.

II. UPFC And IPFC

A. UPFC (Unified Power Flow Controller)

The Unified Power Flow Controller is comprised of a STATCOM and a SSSC coupled via a Common DC link to allow bi-directional flow of active power between the series output terminals of the SSSC and the shunt output terminals of the STATCOM. Each converter can independently generate (or) absorb reactive power at its own AC terminal. The two converters are operated from a DC link provided by a DC storage capacitor. The configuration of a UPFC is shown in Figure 3.

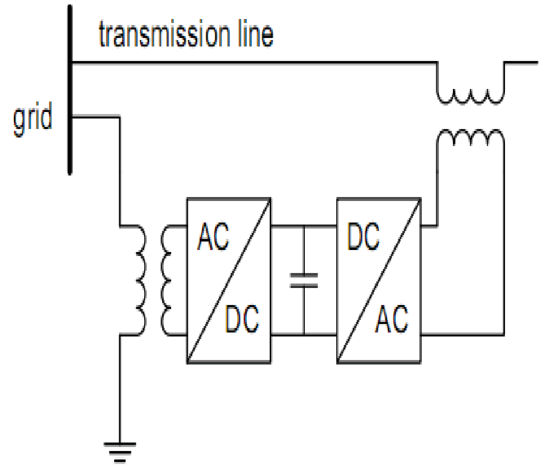


Figure 3. UPFC Configuration

The series converter executes the main function of the UPFC by injecting a voltage, with controllable magnitude and phase angle, in series with the transmission line. It is controlled to provide concurrent active and reactive series compensation without an external energy source. By means of the series voltage injection without angular constraint, the UPFC is able to control, concurrently or selectively, the transmission angle, impedance and line voltage or, alternatively, active and reactive power flow through the line. The link voltage injected by the series converter results in active and reactive power exchange between the series converter and the transmission line. The reactive power is generated internally by the series converter (like SSSC), and the active power is supplied by the shunt converter that is transported through the common DC. The basic function of the shunt converter is to supply or absorb the active power demanded by the series converter. The shunt converter controls the voltage of the DC capacitor by absorbing or generating active power from the bus, therefore it acts as a synchronous source in parallel with the system. Similar to the STATCOM, the shunt converter can also provide independently controllable reactive compensation for the bus.

B.IPFC (Interline Power Flow Controller)

The Interline Power Flow Controller (IPFC) consists of the two (or more) series converters in different transmission lines that are inter-connected via a common DC link, as shown in Figure 4. Unlike other FACTS devices that aim to control the parameter of a single transmission line, the IPFC is conceived for the compensation and control of power flow in a multi-line transmission system [4].

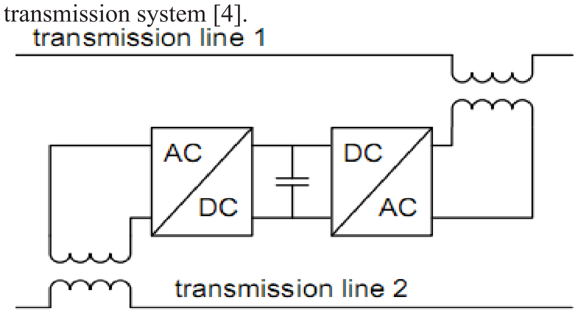


Figure 4. IPFC Configuration

Each converter can provide series reactive compensation of its own line, just as an SSSC can. As the converters can exchange active power through their common DC link, the IPFC can also provide active compensation. This allows the IPFC to provide both active and reactive compensation for some of the lines and thereby optimize the utilization of the overall transmission system. Note that the active power supplied to one line is taken from the other lines. If required, the IPFC can be complemented with an additional shunt converter to provide active power from a suitable shunt bus.

However, the UPFC and IPFC are not widely applied in practice, due to their high cost and the susceptibility to failures. Generally, the reliability can be improved by reducing the number of components; however, this is not possible due to the complex topology of the UPFC and IPFC.

III. Distributed Power Flow Controller(DPFC)

To reduce the failure rate of the components by selecting components with higher ratings than necessary or employing redundancy at the component or system levels are also options. Unfortunately, these solutions increase the initial investment necessary, negating any cost-related advantages. Accordingly, new approaches are needed in order to increase reliability and reduce cost of the UPFC and IPFC at the same time.

The elimination of the common DC link also allows the DSSC concept to be applied to series converters. In that case, the reliability of the new device is further improved due to the redundancy provided by the distributed series converters.

Unlike in UPFC where the active power transfer is through the DC link between the series and shunt converters [7] here in DPFC this power flow is through the transmission lines at the third harmonic frequency which is a zero-sequence component and can be naturally blocked by a Y-Δ transformer. The DPFC makes use of the distributed FACTS (D-FACTS) in the design of the series converter, which is to use multiple single-phase converters instead of one large rated three phase converter while the shunt converter remains as static synchronous compensator (STATCOM) as in UPFC. These large numbers of series converters provides redundancy, thereby increasing the system reliability. As the D-FACTS converters are single-phase and floating with respect to the ground, there is no high-voltage isolation required between the phases. Accordingly, the cost of the DPFC system is lower than the UPFC. The controllability of the DPFC is same as that of the UPFC which refers to the adjustment of the line impedance, the transmission angle, and the bus voltage. The operation principle, the modeling and control, and experimental demonstrations of the DPFC are presented in this paper and the corresponding laboratory results that are carried out on a scaled prototype are also shown.

IV.DPFC Operating Principle And Control

A.DPFC Operating Principle

Within the DPFC, the transmission line presents a common connection between the AC ports of the shunt and the series converters. Therefore, it is possible to exchange active power through the AC ports. The method is based on power theory of non-sinusoidal components.

According to the Fourier Analysis, non- sinusoidal voltage and current be expressed as the sum of sinusoidal functions in different frequencies with different amplitudes. The active power resulting from this non-sinusoidal voltage and current is defined as the mean value of the product of voltage and current. Since the integrals of all the cross product of terms with different frequencies are zero, the active power can be expressed

$$P = \sum_{i=1}^{\infty} V_i I_i \cos \phi_i \qquad (3.1)$$

Where V_i and I_i are the voltage and current at the i^{th} harmonic frequency respectively, and φ_i is the corresponding angle between the voltage and current. Equation (3.1) shows that the active powers at different frequencies are independent from each other and the voltage or current at one frequency has no influence on the active power at other frequencies. The independence of the active power at different frequencies gives the possibility that a converter without a power source can generate active power at one frequency and absorb this power from other frequencies.

By applying this method to the DPFC, the shunt converter can absorb active power from the grid at the fundamental frequency and inject the power back at a harmonic frequency. This harmonic active power flows through a transmission line equipped with series converters. According to the amount of required active power at the fundamental frequency, the DPFC series converters generate a voltage at the harmonic frequency, thereby absorbing the active power from harmonic components. Neglecting losses, the active power generated at the fundamental frequency is equal to the power absorbed at the harmonic frequency. For a better understanding, Figure 5 indicates how the active power is exchanged between the shunt and the series converters in the DPFC system.

The high-pass filter within the DPFC blocks the fundamental frequency components and allows the harmonic components to pass, thereby providing a return path for the harmonic components. The shunt and series converters, the high pass filter and the ground form a closed loop for the harmonic current.

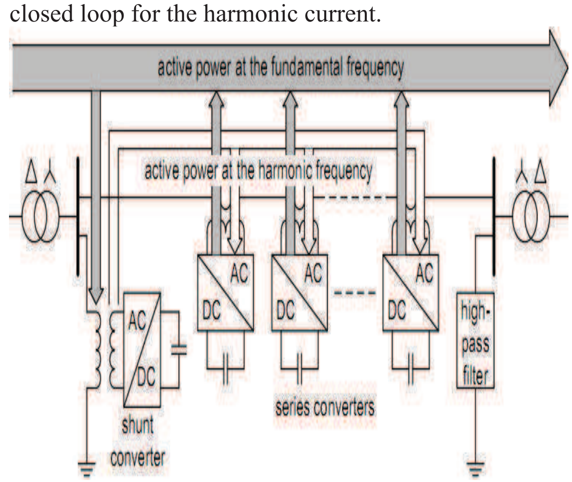


Figure 5. Active power exchange between DPFC converters

B.DPFC Control

To control multiple converters, a DPFC consists of three types of controllers: central control, shunt control and series control, as shown in Figure 5.

The shunt and series control are localized controllers and are responsible for maintaining their own converters' parameters. The central control takes care of the DPFC functions at the power system level. The function of each controller is listed below.

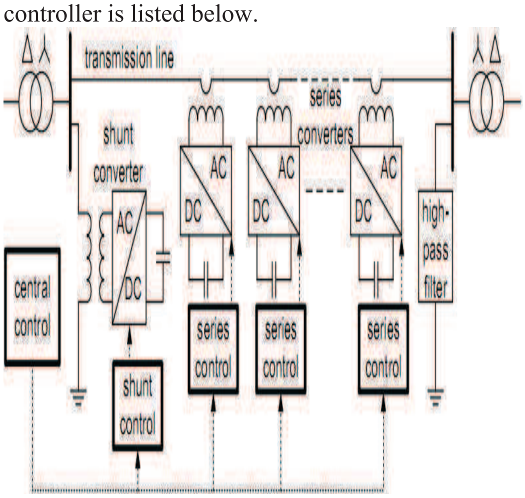


Figure 5. DPFC control block diagram

(a) Central control: The central control generates the reference signals for both the shunt and series converters of the DPFC. Its control function depends on the specifics of the DPFC application at the power system level, such as power flow control, low frequency power oscillation damping and balancing of asymmetrical components. According to the system requirements, the central control gives corresponding voltage reference signals for the series converters and reactive current signal for the shunt converter. All the reference signals generated by the central control concern the fundamental frequency components.

(b) Series control: Each series converter has its own series control. The controller is used to maintain the capacitor DC voltage of its own converter, by using 3rd harmonic frequency components, in addition to generating series voltage at the fundamental frequency as required by the central control.

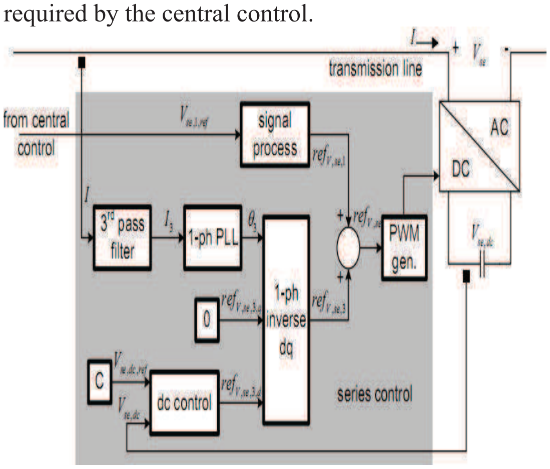


Figure 6. Control scheme of the series converter

(c) Shunt control: The objective of the shunt control is to inject a constant 3rd harmonic current into the line to supply active power for the series converters. At the same time, it maintains the capacitor DC voltage of the shunt converter at a constant value by absorbing active power from the grid at the fundamental frequency and injecting the required reactive current at the fundamental frequency into the grid.

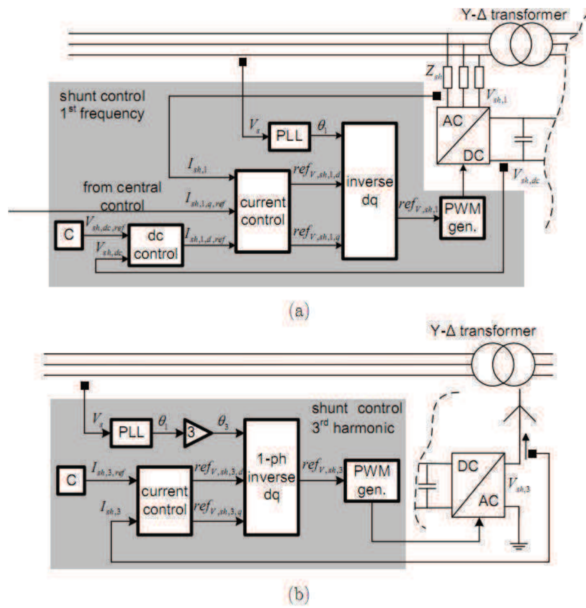
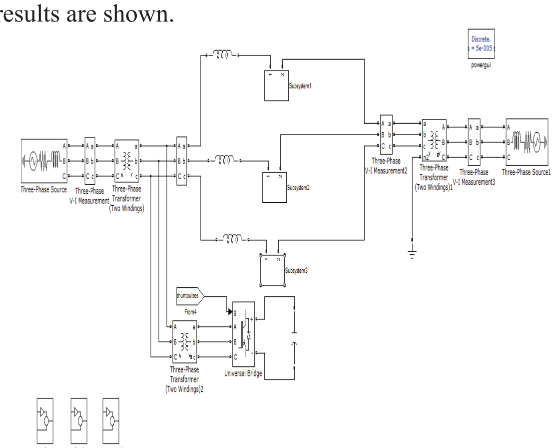


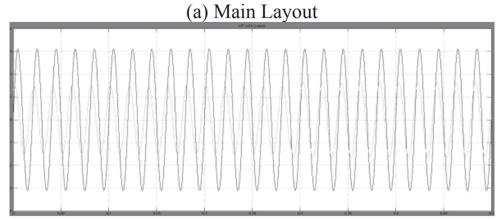
Figure 7. **Control scheme of the shunt converter:**
(a) for the fundamental frequency components
(b) for the 3rd harmonic frequency components

V. Simulation Results

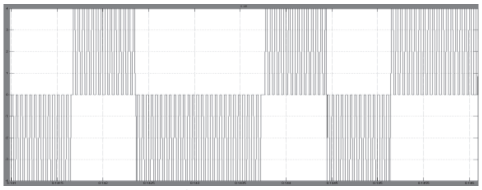
The DPFC model is developed in the MATLAB Simulink environment. The main layout and simulation results are shown.



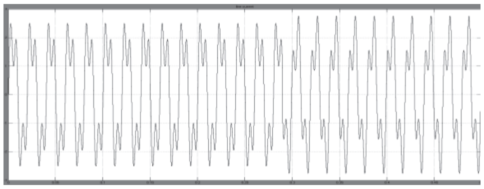
(a) Main Layout



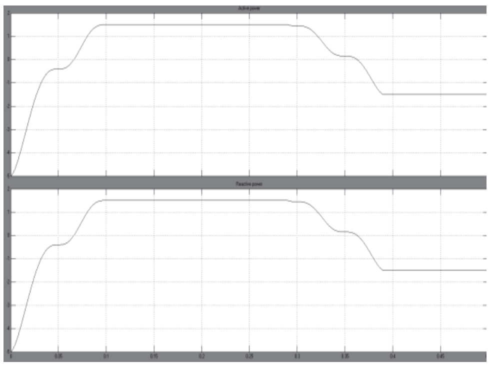
(b) Input Voltage(Vs) Input Current



(c) Series Converter Output



(d) Line Current



(e) Active Power and Reactive Power

CONCLUSIONS

The DPFC basic control is developed based on the dynamic model. The basic control stabilizes the level of the capacitor DC voltage of each converter and ensures that the converters inject the voltages into the network according to the command from the central control. The shunt converter injects a constant current at the 3rd harmonic frequency, while its DC voltage is stabilized by the fundamental frequency component. For the series converter, the reference of the output voltage at the fundamental frequency is obtained from the central control and the DC voltage level is maintained by the 3rd harmonic components. The control parameters of the basic control are determined. Both the model and the basic control are verified in Matlab Simulink.

Utilizing the DPFC to damp low-frequency power oscillation is investigated. The DPFC is used to damp the inter-area oscillatory modes. Because the DPFC can simultaneously adjust three system parameters, namely the bus voltage, the line impedance and the transmission angle, a maximum of three POD controllers can be applied to one DPFC. Within the thesis, the POD controller is designed using the residue method and a two-area network is used in the case study. From the simulation, it can be seen that the DPFC can shift three critical oscillatory modes at the same time. Therefore, it can be concluded that the DPFC has the same capability as the UPFC for power oscillation damping.

Employing the DPFC for asymmetry compensation is studied. Because of the active power exchange between the shunt and the series converters, the DPFC can compensate both active and reactive asymmetry at the fundamental frequency. In addition, since the series converter is single-phase converter, the DPFC can compensate for both zero and negative sequence components. Accordingly, the DPFC currently is the most versatile device for asymmetry compensation among all FACTS devices.

REFERENCES

[1] L. Gyugyi, C. D. Schauder, S. L. Williams, T. R. Rietman, D. R. Torgerson, and A. Edris. “The unified power flow controller: a new approach to power transmission control”. Power Delivery, IEEE Transactions on, 1995.

[2] L. Gyugyi, K. K. Sen, and C. D. Schauder. “The interline power flow con-troller concept: a new approach to power flow management in transmission systems”. Power Delivery, IEEE Transactions on, 1999.

[3] D. Van Hertem, J. Verboomen, R. Belmans, and W. L. Kling. “Power flow controlling devices: an overview of their working principles and their applica-tion range”. In: Future Power Systems, International Conference on, 2005.

[4] R. L. Vasquez Arnez. “The Interline Power Flow Controller: Further aspects related to its operation and main limitations”. In: Transmission andDistribution Conference and Exposition, IEEE/PES, 2008.

[5] Yuan, Z.; de Haan, S.W.H.; Ferreira, B.: “UPFC with Eliminated Common DCLink Connection between Shunt and Series Part”, IEEE Power Engineering Society General Meeting (PESGM) 2007, Tampa, USA.

[6] Yuan, Z.; de Haan, S.W.H.; Ferreira, B.: “A New FACTS Component – Distributed Power Flow Controller (DPFC)”, European Conference on Power Electronics Applications (EPE) 2007, Aalborg, Denmark.

[7] Yuan, Z.; de Haan, S.W.H.; Ferreira, B.: “A Method of Transmitting Active Power Between Converters with Eliminated Common DC Link”, IEEE Benelux Young Research Symposium in Electrical Power Engineering 2008, Eindhoven, the Netherlands.

Multilevel Inverters Fed PMSM Drive using Carrier Based Space Vector Pulse Width Modulation

G. Sree Lakshmi[1], Dr. S. Kamakshaiah[2] and Dr. G. Tulasi Ram Das[3]
[1] CVR College of Engineering, Department of EEE, Ibrahimpatan, R.R.District, A.P., India
Email: s_sreelakshmi@yahoo.com
[2] JNT University, Department of EEE, Hyderabad, AP, India.
Email: skamakshaiah@yahoo.com
[3] Vice Chancellor, JNT University, Kakinada, A.P., India
Email:das_tulasi@gmail.com

Abstract--- **Space Vector Pulse Width Modulation (SVPWM), one of the advanced computation based PWM techniques, has many advantages over conventional carrier-based PWM methodologies. This paper presents a systematic approach and simplified algorithm of space vector pulse width modulation (SVPWM) for a two-level & three level three-phase inverter, which can operate in under modulation and over modulation modes. This method is called as Carrier Based Space Vector Pulse Width Modulation (CBSVPWM) which is based on description of controllable redundant parameters in the modulating signals. The CBSVPWM methods can be described in a unified mathematical formulation, and obtain the same outputs similarly as of corresponding SVPWM. The simulation and analysis of CBSVPWM to the two and three level inverters fed to Permanent Magnet Synchronous Motor drives and their differences and advantages are also presented.**

Index Terms- **Multilevel Inverters, Space Vector Pulse Width Modulation (SVPWM), Carrier Based Space Vector Pulse Width Modulation (CBSVPWM), Permanent Magnet Synchronous Motor (PMSM).**

I. INTRODUCTION

The smaller voltage steps created by a multilevel converter lead to the production of higher power quality waveforms and also reduce the *dv*/*dt* stresses on the load, especially in AC motors, and reduce the electromagnetic compatibility problems [1]. Permanent magnet synchronous motors (PMSM) are widely used in high-performance drives such as industrial robots and machine tools for their advantages high power density, high-torque, and free maintenance and so on. In recent years, the magnetic and thermal capabilities of the PM have been considerably increased by employing the high-coercive PM materials [2]. In last few years permanent magnet synchronous motor (PMSM), consequently is acquired in more and more far-ranging application, because of its properties such as small volume, light weight, high efficiency, small inertia, rotor without heat problem, etc. [3].

A variable voltage can be obtained by varying the input DC voltage and maintaining the gain of the inverter constant. On the other hand if the DC input voltage is fixed and it is not controllable, a variable output voltage can be obtained by varying the gain of the inverter, which is normally accomplished by PWM/SVPWM control with in the inverter. The output voltage waveform of a practical inverters are non sinusoidal and contains certain harmonics. For high power applications low distorted sinusoidal waveforms are required. It is possible to achieve high-voltage low-distortion ac waveforms by increasing the number of dc levels that are available in the inverter. Thus the concept of Multilevel Inverters is introduced. Multilevel inverters are an array of power semiconductors and capacitor voltage sources, the output of which generate voltages with stepped waveforms [4]. Using enough levels the multi level inverter generates approximately a sinusoidal voltage waveform with very low harmonic distortion. There are three types of multilevel inverters as NPC multilevel inverters, cascade multilevel inverter and dc capacitor clamped multilevel inverter. The most common configuration of multilevel inverters is the neutral point clamped voltage source inverter structure (VSI-NPC) which is widely used in medium voltage drives for rolling mills, marine, and traction applications.

To control multilevel converters, the pulse width modulation (PWM) strategies are the most effective, especially the space vector pulse width modulation (SVPWM) one, which has equally divided zero voltage vectors describing a lower total harmonic distortion (THD). By and large, the emphasis has been placed on space vector PWM (SVPWM) methods. SVPWM offers great flexibility to optimize switching waveforms and is suited for digital implementation [5]. However, due to constant sampling rate used in SVPWM and the complications associated with this method when we go to n-level, the equivalent carrier-based techniques have been developed.

Carrier-based space vector modulation (CBSVM) is appropriate for inverters with more than five levels, where the computational overhead for conventional SVPWM is exceeding due to many output states. Wenxi Yao proposed carrier-based space vector modulation technique [8], which are harmonically equivalent, with the best spectral performance being achieved when the nearest three space vector states are selected with the middle two vectors centered in each half carrier switching interval. This strategy is known as carrier based space vector modulation (CBSVM).

The speed or torque of a Permanent Magnet Synchronous Motor can be controlled by various modulation strategies for inverter. In this paper control of PMSM in open loop system using the best modulation strategy known as CBSVM technique is simulated for two and three level inverters and shown that Permanent Magnet Synchronous Motor performance is improved. The three level performance of PMSM is improved when compared to the two level inverter using CBSVM technique. The performance of the PMSM can be further improved.

II. Multilevel Inverters

The demand for high voltage inverter drives and for power transmission has been rising in the recent years. For such applications, standard PWM voltage source inverters with only two switches per phase place requirements on the blocking voltage of the switches that cannot be met with a single device. Currently, to overcome this challenge we are using multilevel converters. Multilevel inverter technology has emerged recently has a very important alternative in the area of high-power medium-voltage energy control.

Multilevel inverters include an array of power semiconductors and capacitor voltage sources, the output of which generate voltages with stepped waveforms. Commutation of the switches permits the addition of the capacitor voltages, which reach high voltage at the output, while the power semiconductors must withstand only reduced voltages. Figure 1.[6] shows a schematic diagram of one phase leg of inverters with different number of levels.

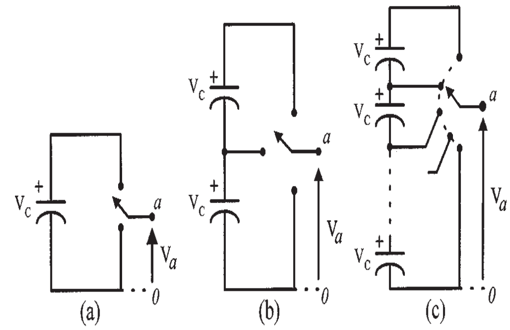


Figure 1. One phase leg of an inverter with (a) two level (b) three level (c) 'n' level

The action of the power semiconductors is represented by an ideal switch with several positions. A two-level inverter generates an output voltage with two values (levels) with respect to the negative terminal of the capacitor while the three-level inverter generates three voltages, and so on.

The term multilevel starts with the three-level inverter. By increasing the number of levels in the inverter, the output voltages have more steps generating a staircase waveform, which has a reduced harmonic distortion. However, a high number of levels increases the control complexity and introduces voltage imbalance problems.

III. Carrier-Based SVPWM

Carrier based SVPWM allow fast and efficient implementation of SVPWM without sector determination. The technique is based on the duty ratio profiles that SVPWM exhibits. By comparing the duty ratio profile with a higher frequency triangular carrier the pulses can be generated, based on the same arguments as the sinusoidal pulse width modulation [8].

The power circuit of a three-phase voltage-source inverter (VSI) is shown in Figure 2., V_a, V_b, and V_c are the three output voltages applied to the star-connected motor windings, and where V_{dc} is the continuous inverter input voltage. This is a two-level inverter consisting of six power transistors S1 through S6, which are controlled by switching signals.

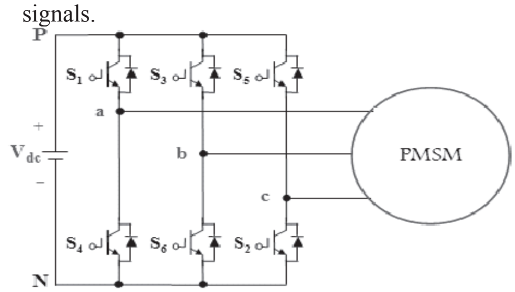


Figure 2. Three-phase VSI bridge circuit

When an upper transistor is switched on, the corresponding lower transistor is switched off. There are eight different combinations (000), (100), (110), (010), (011), (001), (101), and (111). The first and last states do not cause a current to flow to the motor, and hence, the line-to-line voltages are zero. The other six states can produce voltages to be applied to the motor terminals. If the inverter operation starts by state (100) to be state 1, it is possible to compute the voltage space vectors for all inverter states which are shown in the complex space vector plane in Fig.3 [3]. The six active voltage space vectors are of equal magnitude (2/3) V_{dc} and mutually phase displaced by 60^o, as shown in Figure 3.

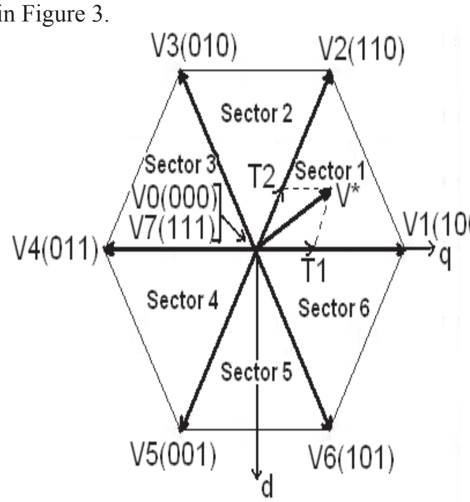


Figure 3. Voltage space vectors for a three-phase VSI

Note that the switching states of each arm should be combined with each other to compose the required three-phase output voltage. Because each pole voltage has only two levels according to the related switching state, the time duration in which the different voltages are maintained is definitely related to the voltage modulation task. Therefore, the modulation task can be greatly simplified by considering the relation between the time duration and the output voltage[9]. We now focus on the effective voltage that makes an actual power flow between inverter and load. Fig.4 shows the switching states of sector 1 at different times during two sampling intervals. T_S denotes the sampling time and T_{eff} denotes the time duration in which the different voltage is maintained. T_{eff} is called the "effective time". For the purpose of explanation, an imaginary time value will be introduced as follows:

$$Txs = \frac{Ts}{Vdc} Vxs^* \text{ ,(x=a,b,c)}$$

V_{as}*,V_{bs}*and V_{cs}* are the A-phase, B -phase, and C-phase reference voltages, respectively. This switching time could be negative in the case where negative phase voltage is commanded. Therefore, this time is called the "imaginary switching time".

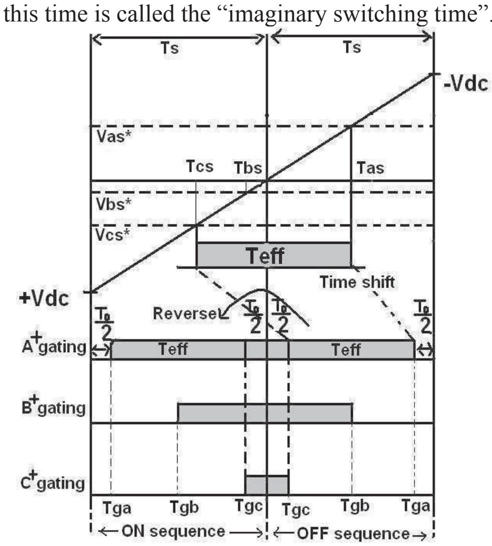


Figure 4. Actual gating time generation for continuous SVM

Now, the effective time can be defined as the time duration between the minimum and the maximum value of three imaginary times, as given by

$$T_{eff} = T_{max} - T_{min}$$

Where $T_{min} = \min(T_{as},T_{bs},T_{cs})$

$$T_{max} = \max(T_{as},T_{bs},T_{cs})$$

When the actual gating signals for power devices are generated in the PWM algorithm, there is one degree of freedom by which the effective time can be relocated anywhere within the sampling interval.

Therefore, a time-shifting operation will be applied to the imaginary switching times to generate the actual gating times (T_{ga},T_{gb},T_{gc}) for each inverter arm, as shown in Fig. 4. This task is accomplished by adding the same value to the imaginary times as follows:

$$T_{ga}=T_{as}+T_{offset}$$
$$T_{gb}=T_{bs}+T_{offset}$$
$$T_{gc}=T_{cs}+T_{offset}$$

Where T_{offset} is the 'offset time'.

This gating time determination task is only performed for the sampling interval in which all of the switching states of each arm go to 0 from 1. This interval is called the "OFF sequence". In the other sequence, it is called the "ON sequence." In order to generate a symmetrical switching pulse pattern within two sampling intervals, the actual switching time will be replaced by the subtraction value, with sampling time as follows:

$$T_{ga}=T_s-T_{ga}$$
$$T_{gb}=T_s-T_{gb}$$
$$T_{gc}=T_s-T_{gc}$$

IV. MODELLING OF PMSM

The stators of the PMSM and the wound rotor synchronous motor (SM) are similar and there is no difference between the back EMF produced by a permanent magnet and that back EMF produced by an excited coil. Hence the mathematical model of a PMSM is similar to that of the wound rotor SM [11], [12].

The stator d, q equations of the PMSM in the rotor reference frame are:

$$v_q = ri_q + P\omega_r\lambda_d + p\lambda_q$$

$$v_d = ri_d - P\omega_r\lambda_q + p\lambda_d$$

Where

$$\lambda_q = L_q i_q$$

$$\lambda_d = L_d i_d + \lambda_m$$

Where P is the pole pairs, p is the d/dt operator, v_q and v_d are the q, d axis voltages, i_q and i_d are the q, d axis stator currents, L_q and L_d are the q, d axis inductances, λ_q and λ_d are the q, d axis stator flux linkages, while r and ω_r are the stator resistance and rotor speed, respectively. λ_m is the flux linkage due to the rotor magnets linking the stator.

The electromechanical torque developed by the motor is:

$$T_{em} = \frac{3}{2}P(\lambda_d i_q - \lambda_q i_d)$$

By substituting the values of λ_q and λ_d in the above equation

$$T_{em} = \frac{3}{2}P(\lambda_m i_q + (L_d - L_q)i_q i_d)$$

The relationship between the electromechanical torque and the load torque is given as:

$$\frac{d\omega_r}{dt} = \frac{1}{J_m}(T_{em} - T_l - B_m\omega_r)$$

$$\frac{d\theta_r}{dt} = \omega_r$$

For dynamic simulation, the equations of PMSM presented above must be expressed in state-space form as the following:

$$pi_q = \frac{1}{L_q}(v_q - ri_q - P\omega_r L_d i_d - P\omega_r\lambda_m)$$

$$pi_d = \frac{1}{L_d}(v_d - ri_d + P\omega_r L_q i_q)$$

$$p\omega_r = \frac{1}{J_m}(T_{em} - T_l - B_m\omega_r)$$

V. SIMULATION AND RESULTS

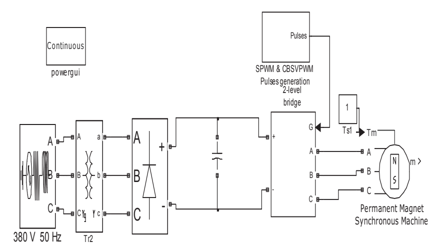


Figure5. Simulink model for three phase two level inverter fed PMSM Drive

A. Simulation Results of two level inverter fed PMSM drive using CBSVPWM

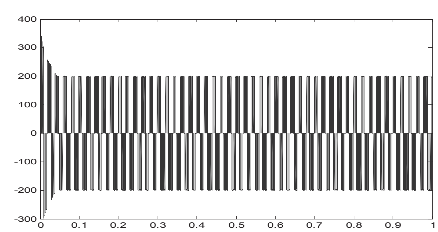


Figure 6.Output Phase Voltage of a two level inverter

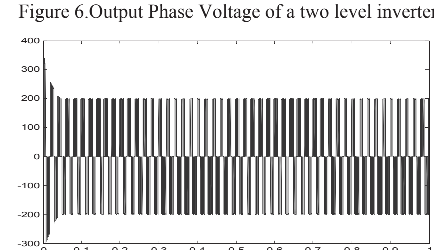


Figure 7.Output Line Voltage of a two level inverter

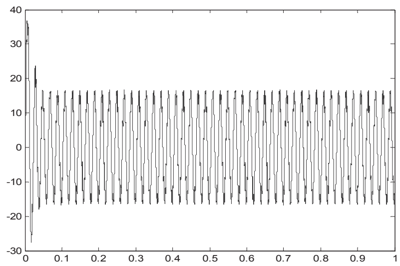


Figure 8.Output Line current of a two level inverter

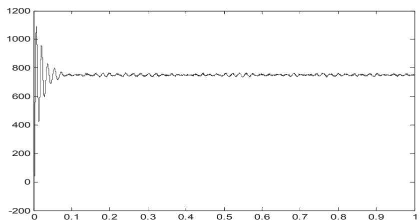


Figure 9.Output Speed of a two level inverter

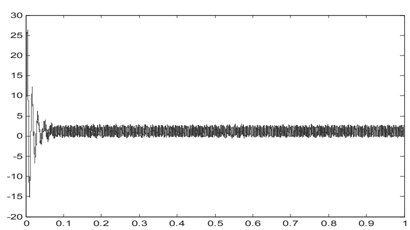


Figure 10.Output Torque of a two level inverter

B. Simulink model of three-level inverter fed PMSM Drive using CBSVPWM:

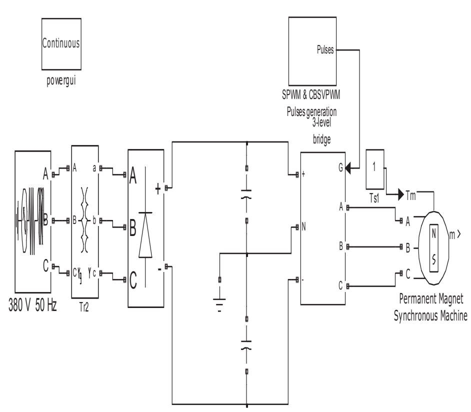


Figure 11.Simulink model for three phase three level inverter fed PMSM Drive

C. Simulation Results of three level inverter fed PMSM drive using CBSVPWM

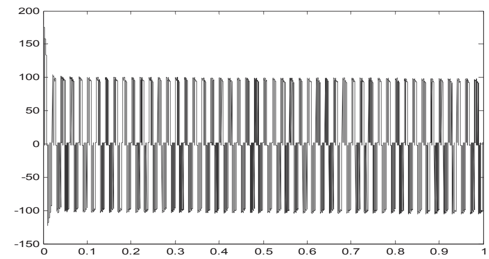


Figure 13.Output Phase Voltage of a three level inverter

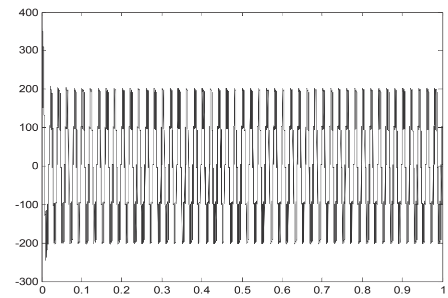


Figure 14.Output Line Voltage of a three level inverter

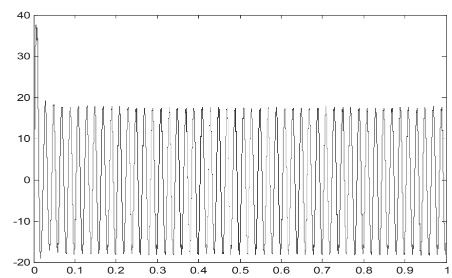


Figure 15.Output Line current of a three level inverter

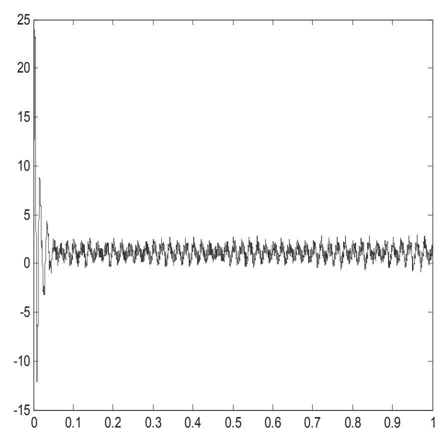


Figure 16.Output Torque of a three level inverter

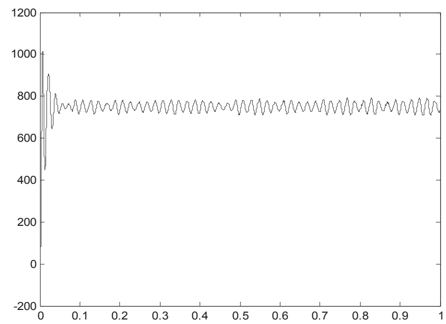


Figure 17.Output Speed of a three level inverter

Determination of harmonic distortion: By performing FFT analysis we can determine the harmonic distortion in the line current and line voltages.

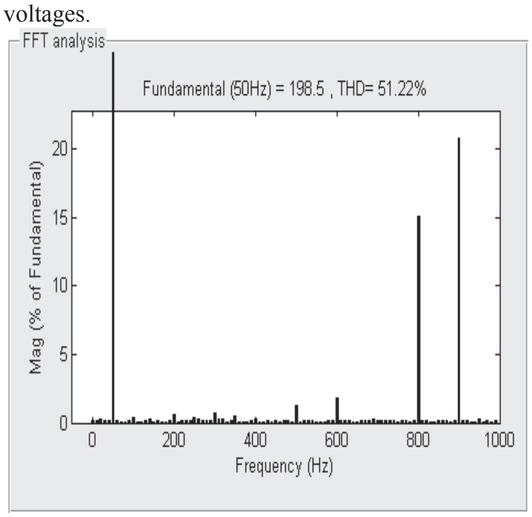


Figure 18.Determination of THD of line voltage of two level inverter

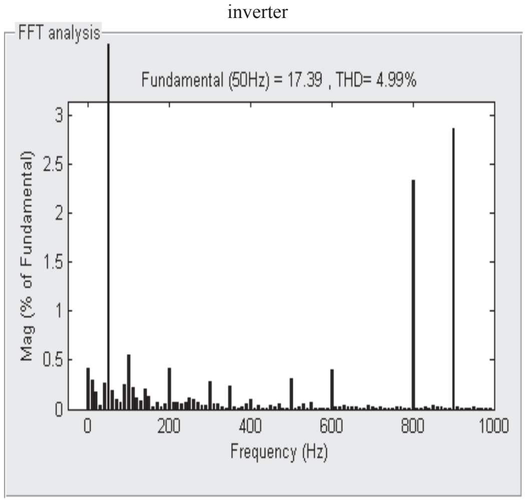


Figure 19.Determination of THD of line current of two level inverter

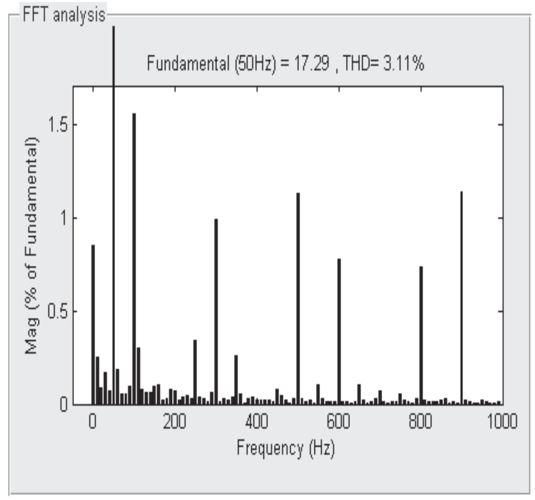


Figure 20.Determination of THD of line current of three level inverter

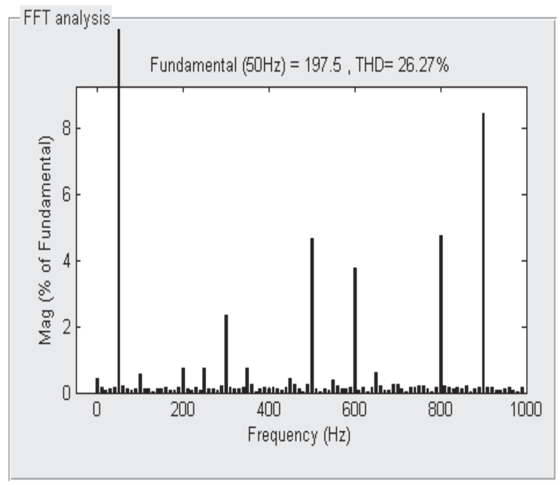


Figure 21.Determination of THD of line voltage of three level inverter

CONCLUSION

Table I.
Comparison of THD of Two and Three level inverters fed PMSM Drive using CBSVPWM

THD	Two level inverter	Three level inverter
Line voltage	51.22	26.27
Line current	4.99	3.11

In this paper, simulation analyses concerning the applications of CBSVPWM control strategy on the two and three level inverters fed Permanent Magnet Synchronous Motor are presented. Carrier Based Space Vector Pulse Width Modulation gives the same results as that of Space Vector Pulse width Modulation, more over it reduces the calculation of dwelling times which simplifies the method .From this analysis author can conclude that multilevel inverter can eliminate the harmonics produced by the normal inverter. From the simulation results obtained we can say that the total harmonic distortion reduces by increasing the number of levels in the output voltage. The switching losses are also reduced by using this technique.

The THD from Table1, for the three level and two level are obtained and observed that the three level produces less harmonics and better speed, torque characteristics compared to two level. Therefore as the level increases we can further decrease the harmonic content and can improve the performance of the drive. Not only this CBSVPWM can easily implemented for n-level inverters.

REFERENCES

[1] B. Wu, *High Power Converters and AC Drives*, Wiley- IEEE Press, Hoboken, New Jersey, 2006.

[2] I. Takahashi and T. Naguchi, A new quick response and high-efficiency control strategy of an induction motor, IEEE Trans. Ind. Appl., Vol IA-22, pp. 820-827, 1986.

[3] Li Ye and Yan Xinpin, "The perspective and status of PMSM electrical servo system,"Micromotors Servo Technique, 2001, Vol.4, pp. 30-33.

[4] "Multilevel converters-a new breed of power converters" by Jih-Sheng Lai, and Fang Zheng Peng.

[5] "Multilevel inverters: a survey of topologies, controls, and applications" by José Rodríguez, Jih-Sheng Lai, and Fang Zheng Peng 1999 IEEE paper.

[6] "Multilevel converters for large electric drives" by Leon M. Tolbert, Fang Zheng Peng, and Thomas G. Habetler.

[7] "A simple svpwm algorithm for multilevel inverters" by Lei Hu, Honyan Wang, Yan Deng and Xiangning He College of Electrical Engineering, Zhejiang University Hang Zhou, China. 2004 35th Annual IEEE Power Electronics Specialists Conference.

[8] "A Comprehensive Analysis of Space Vestor PWM Technique based on placement of zero-space vector"by G.Sambasiva Rao & Dr.K.Chandra Sekhar.

[9] "Unified Voltage Modulation Technique for Real Time Three-phase Power Conversion" by Dae-Woong Chung, Joohn-Sheok Kim and Seung-Ki Sul.

[10] "Control of Voltage Source Inverters using PWM/SVPWM for Adjustable Speed Drive Applications" by Mr. Anupam Mishra, Mr. Sunil Panda, Mr. B. Srinivas.

[11] P. Pillay , and R. Krishnan, "Modeling, simulation, and analysis of permanent-magnet motor drives, part I: the permanent-magnet synchronous motor drive," *IEEE Trans. Industry Applications,* vol. 25, no. 2, March/April. 1989, pp. 265-273.

[12] P. Pillay , and R. Krishnan, "Modeling of permanent magnet motor drives," *IEEE Trans. Industrial Electronics,* vol. 35, no. 4, Nov. 1988, pp. 537-541.

Multi-Vendor Remote Monitoring and Control System for Hydraulic Drive Systems

Naregalkar Akshay
CVR College of Engineering, Department of ECE, Ibrahimpatan, R.R.District, A.P., India
Email: akshay.naregalkar@gmail.com

***Abstract*— The main idea of this paper is to interface two most powerful technologies LabVIEW (Laboratory Virtual Instrument Engineering Workbench) ,PLC(Programmable Logic Controller) and internet control which are ruling the industries and lead to many new features like acquiring data at faster rate, controlling the process accurately, providing multi-tasking operations, remote controlling and several other features. The aim of this paper is to provide more effective and efficient process control and automation by interfacing Allen-Bradley PLC with LabVIEW 10.0 software using OPC Server which is a complete communication server providing plant and floor device connectivity for a wide variety of platforms. Also to control the industrial processes remotely with web publishing tool which can implement a powerful real-time performance monitoring and control system using graphical programming with LabVIEW. Thereby creating a multivendor system, in Automation Company which can choose and apply the best products and systems for any given application often requires investing a significant amount of money on system integration. A Hydraulic Drive System is connected to PLC I/O module and controlled remotely through Internet using OPC as an interface between PLC and LabVIEW.**

***Index Terms*—LabVIEW, PLC, web publishing, remote monitoring, control, automation**

I. INTRODUCTION

Most manufacturers pressing need of making hardware and software work together. The main problem of manufacturers is interfaces not standard. The proprietary system not integrates among each other. The solution is having a standard that provides real plug-and-play software technology for process control and factory automation for every system, every device and every driver can freely communicate, connect and integrated. The standard is OPC (OLE for Process Control) where OLE is Object Linking and Embedding. OPC will mean the dawning of a new day for users of industrial software and hardware.

Also remotely control applications over a wide area had been commonly used in the industries today. One of the common applications requires remote control and monitoring is Stepper motor drive system. Drive system has various types of controller, in order to perform some actions such as control the speed, forward and reverse turning direction of the motor. This approach can be done by Programmable Logic Controller (PLC), and with the rise of the technology, will be used in order to achieve the remote control system. Plus the PLC today can be controlled not only using its original software, but 3rd party software as well, such as LabVIEW. Whereas the OPC Server will provide interface between LabVIEW and PLC.

With this en-user will receive benefit not only from the improvement in salability and integration, but also from continual improvement in technology. Vendors of industrial application, who until now spent time and money on developing software incompatible with other vendors' products, will now make an effort to develop high-quality products totally compatible with all applications. Improvement in function, quality and service will be possible throughout the industry. And enabling the remote panels feature of LabVIEW in a process that walks the user through the creation of a Web page that automatically embeds the appropriate LabVIEW application into the new Web Page.

This paper implements the idea of adding advanced analysis and control functionality to any PLC using the OPC servers with LabVIEW software and web based control for real time systems to solve the problem of remote monitoring and control.

II. OPC SERVER FOR LABVIEW – PLC INTEGRATION

A Programmable Logic Controller (PLC) is a digital computer used for automation of processes which is designed for multiple inputs and outputs. Programs to control machine operation are typically stored in battery-backed or non-volatile memory. A PLC is an example of a hard real time system since output results must be produced in response to input conditions within a bounded time, otherwise unintended operation will result.

But one major disadvantage of PLCs is lack of standardization. This causes a lot of confusion if the PLC used for an application is replaced by one from a different manufacturer, or if a PLC programmer is replaced by a person with a different understanding of PLC programming. The Ladder Logic or FBD Logic which are used to program PLCs are very tedious to work with. This makes the programming a very challenging and difficult job.

LabVIEW ties the creation of user interfaces (called front panels) into the development cycle. LabVIEW programs/subroutines are called virtual instruments (VIs). Each VI has three components: a block diagram, a front panel, and a connector panel. The last is used to represent the VI in the block diagrams of other, calling VIs. Controls and indicators on the front panel allow an operator to input data into or extract data from a running virtual instrument. However, the front panel can also serve as a programmatic interface. Thus a virtual instrument can either be run as a program, with the front panel serving as a user interface, or, when dropped as a node onto the block diagram, the front panel defines the inputs and outputs for the given node through the connector pane. This implies each VI can be easily tested before being embedded as a subroutine into a larger program.

The graphical approach also allows non-programmers to build programs simply by dragging and dropping virtual representations of lab equipment with which they are already familiar. With LabVIEW Creating our application is as simple as dragging and dropping graphical functions and wiring the objects together to form a dataflow program with LabVIEW which is a graphical programming. Before OPC, we had to write separate client application code to communicate with each device.

OLE for Process Control (OPC), which stands for Object Linking and Embedding (OLE) for Process Control, is the original name for a standard specification developed in 1996. OPC servers provide a method for many different software packages to access data from a process control device, such as a PLC (Programmable Logic Controller) or DCS (Distributed Control System). Traditionally, any time a package needed access to data from a device, a custom interface, or driver, had to be written. The purpose of OPC is to define a common interface that is written once and then reused by any business, SCADA, HMI, or custom software packages.

Once an OPC server is written for a particular device, it can be reused by any application that is able to act as an OPC client. OPC servers use Microsoft's OLE technology (also known as the Component Object Model, or COM) to communicate with clients. COM technology permits a standard for real-time information exchange between software applications and process hardware to be defined.

The basic concept in OPC is that we have an OPC Server and one or more OPC Clients that communicate with the server in order to write or read data. An OPC server has implemented a set of services, and the clients are using these services. At a high level, an OPC server is comprised of several objects: the server, the group, and the item. The OPC server object maintains information about the server and serves as a container for OPC group objects. The OPC group object maintains information about itself and provides the mechanism for containing and logically organizing OPC items.

An OPC Client can connect to OPC Servers provided by one or more vendors. Tags are used a lot in the process industry and are normally assigned to a piece of information. A tag consists of a name describing a single point of information so a process system can consists of hundreds and even thousands of tags. The OPC server has one tag for each measurement points and controller points in the plant and it is the responsibility of the OPC server to get the information from the controllers. This is one of the reasons for the complexity of the servers; they need to have drivers for a lot of controllers and measurement systems.

OPC Servers provides a single consistent interface to communicate with multiple devices, saving you from learning new communication protocols or spending time understanding new applications. The combination of OPC Servers and LabVIEW provides a single platform for delivering high performance measurements and control to both new and existing industrial systems. The OPC servers connect through the OPC client in LabVIEW Data logging and Supervisory Control (DSC) Module to enable you develop a fully fledged HMI/SCADA system with PLCs, PACs and smart sensors.

The OPC Servers provides a single consistent interface to communicate with multiple devices, saving from learning new communication protocols or spending time understanding new applications. The combination of OPC Servers and LabVIEW provides a single platform for delivering high performance measurements and control to both new and existing industrial systems.

LabVIEW software can communicate with any programmable logic controller (PLC) in a variety of ways. OPC Servers are available for virtually all PLCs.

Now, this cross connectivity between LabVIEW and PLC is done through OPC server which results in complete control of PLC hardware under LabVIEW software as shown in figure 1.

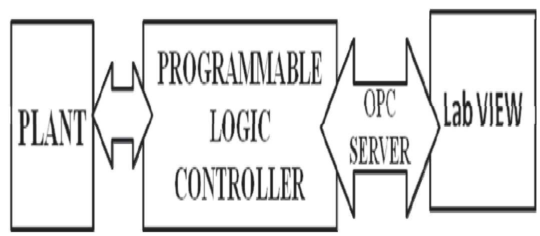


Figure1. PLC LabVIEW Interface with OPC Server

At the other end we use LabVIEW to connect to this OPC server through OPC Client. OPC Client is software that connects OPC servers provided one or more vendors as shown in Figure 2.

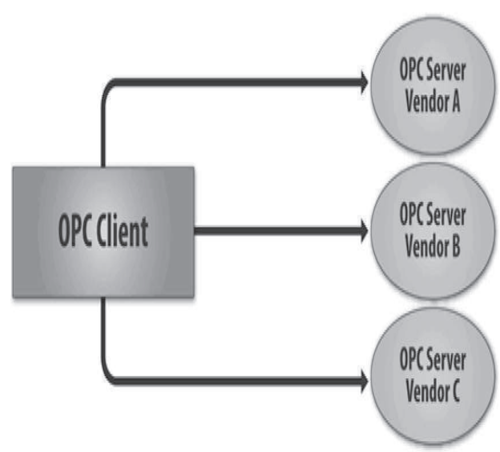


Figure 2. OPC Client and Server

III. HYDRAULIC DRIVE SYSTEM

A Hydraulic Drive System is connected to PLC I/O module and controlled remotely through Internet using OPC as an interface between PLC and LabVIEW. The Hydraulic drive/actuator is used for loading and unloading application as shown in Figure 3.

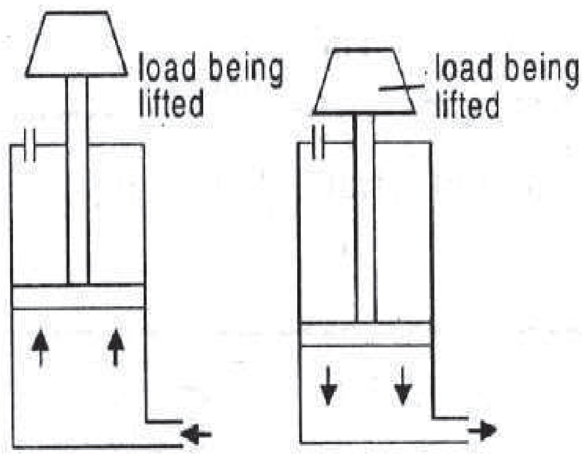


Figure 3. Loading and Unloading with Hydraulic Actuator

Hydraulic systems are used to control & transmit power .A pump driven by prime mover (electric motor) creates flow of fluid. An actuator is used to convert the energy of the fluid back into mechanical power. Amount of output power developed depends upon the flow rate, pressure drop across the actuator & its overall efficiency. Fluid enters through inlet port into piston end or blank end–pressure build up-force generation on piston-movement of piston –extension or forward stroke. Retraction or return by compression spring or under the influence of gravity.

This Hydraulic System is shown in figure 4.

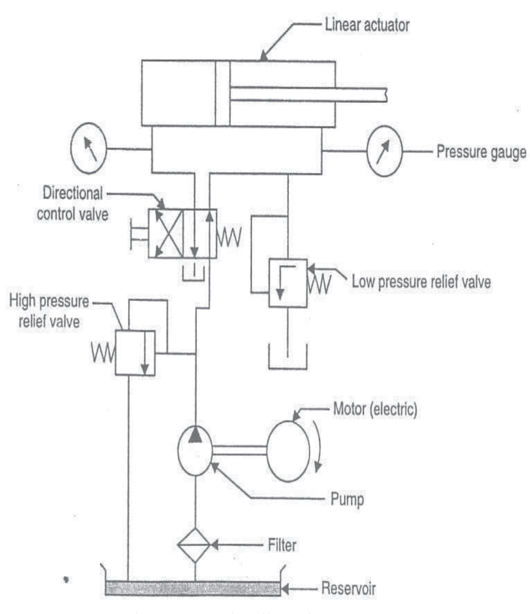


Figure 4. Hydraulic Drive System

Initially the motor is switched ON and then the Hydraulic drive/ Actuator arm is operated based on the requirement and conditions. The condition for activating the Hydraulic drive/ Actuator arm upwards is set when the micro switch sensor (forward) of lower position is active .When the Hydraulic drive/ Actuator arm moves up the upper micro switch sensor is activated and the condition is reset.

The condition for bringing Hydraulic drive/ Actuator arm downwards (retract) is when the micro switch of upper position is active .When the Hydraulic drive/ Actuator arm reaches the lower position, the lower micro switch sensor is activated and the condition is reset. To have forward (up) or Reverse (down) movement of the hydraulic cylinder piston, activate the corresponding solenoids with PLC.

IV. Remote Monitoring And Control With Web Publishing For LabVIEW – Internet Application

Remote Monitoring and Control can be performed in LabVIEW by using the Web Publishing Tool. LabVIEW uses its own web server to publish the front panel to a web browser so the VI can be accessed remotely. Web Publishing Tool is a tool which can help to publish VIs (Virtual Instrument which is a LabVIEW program) on the internet. It generates a URL which can be used to access the VI from any computer with an active internet connection. From the Tools Menu in LabVIEW, select the Web Publishing Tool, The web publishing tool dialog box appears. Change the Viewing Options either Monitor or Embed as shown in figure 5.

Monitor is the simplest kind of remote connection and it allows a remote computer to view the operation of the file. Embed gives a remote connection that allows a remote computer to control the operation of the file. Take this in steps. Click on Save to Disk.

Now save the Webpage to the default location with the default file name. The web publishing tool should then display the web address to use to remotely monitor the VI (program) as shown in figure 6.

Write this web address down as it will be necessary to enter this on a different computer to observe the VI. Then Click OK to close the Document URL dialog box. Then click Start Web Server in the Web Publishing Tool Dialog Box and then click Done. Now Run the program.

One can also select the Preview in Browser button at any time to see what your current settings would look like in a Web browser. The next screen allows to create the document title, header, and footer for the Web page where the VI will be published.

Go to a different computer and open a web browser. Enter the website into the Address box of the web browser, while the program is running. The Front panel should be displayed.

At this point, one can view the user interface over the Web. To obtain control, right-click within the web page and select Request Control of VI. Now the VI can be controlled remotely. One can now run the VI, interact with controls, and view data in indicators as if viewer was running the VI from within the development environment. Another user at a different location can also open a Web Browser and navigate to the specified URL and now have their own instance of this VI that can be controlled without being affected by any other user connected to the URL.

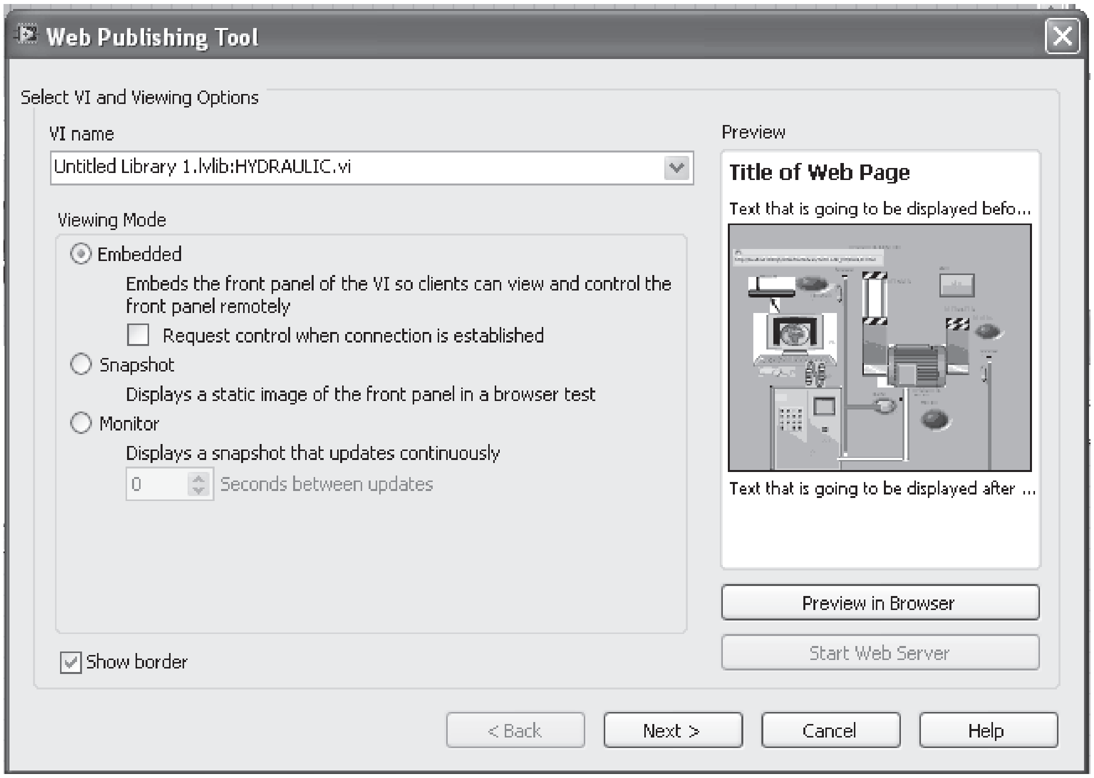


Figure5 .Web Publishing of Remote monitoring and Control of Hydraulic Drive Control

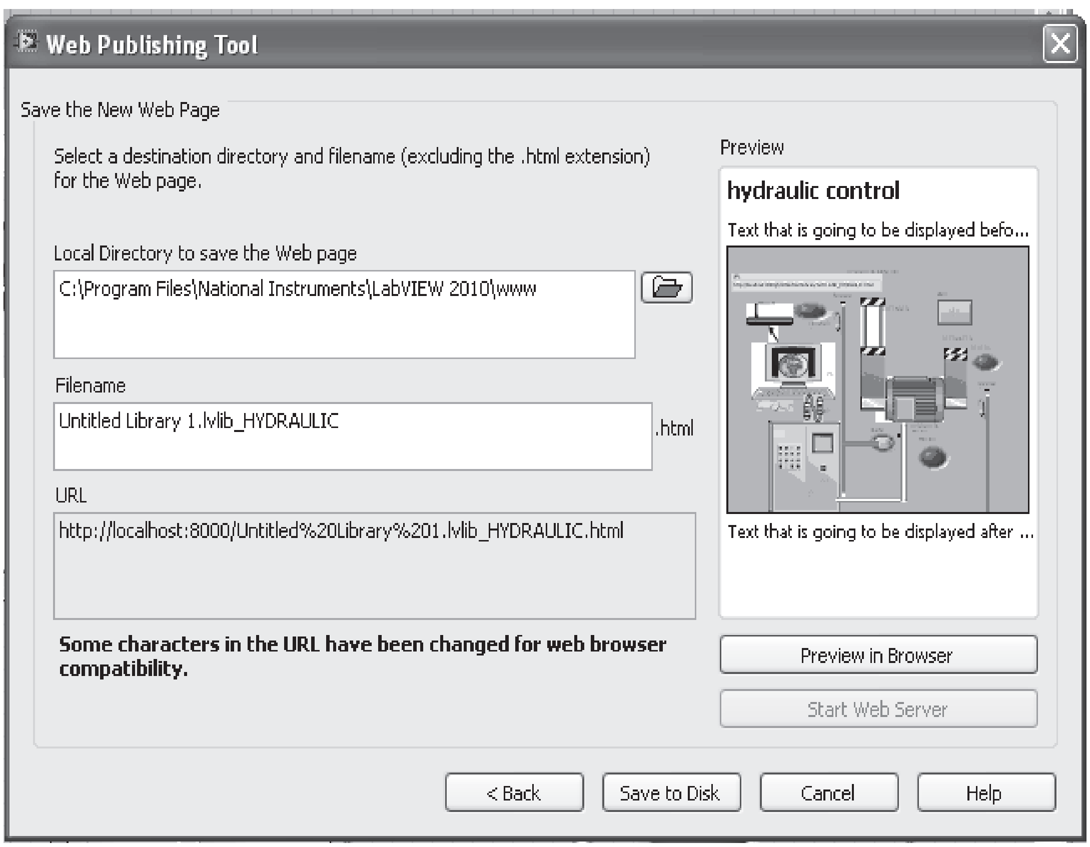


Figure 6.Creating URL for the LabVIEW VI by Web Publishing of Remote monitoring and Control of Hydraulic Drive Control

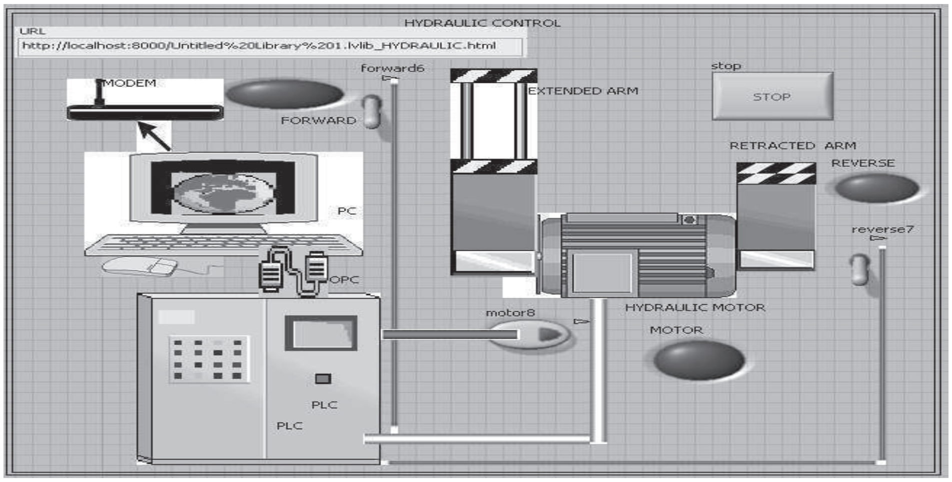


Figure 7. Front Panel of Hydraulic Drive Control with PLC –LabVIEW interface via OPC

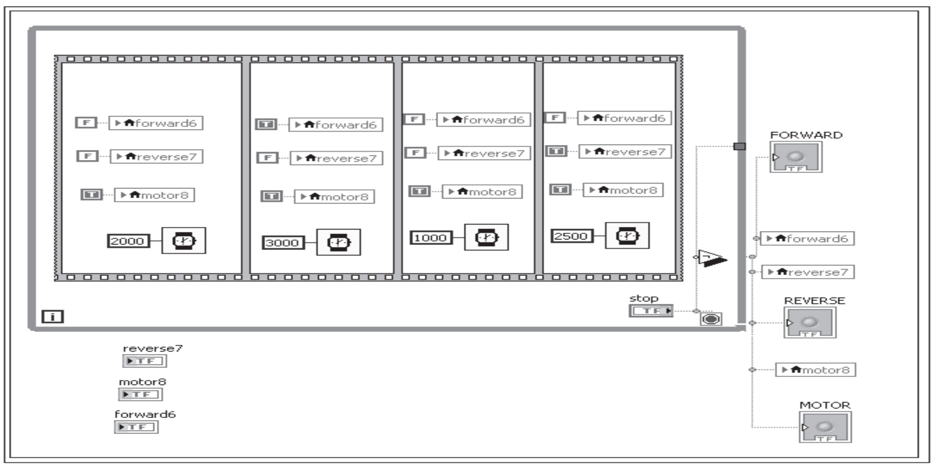


Figure 8. Block Diagram of Hydraulic Drive Control with PLC – LabVIEW interface via OPC

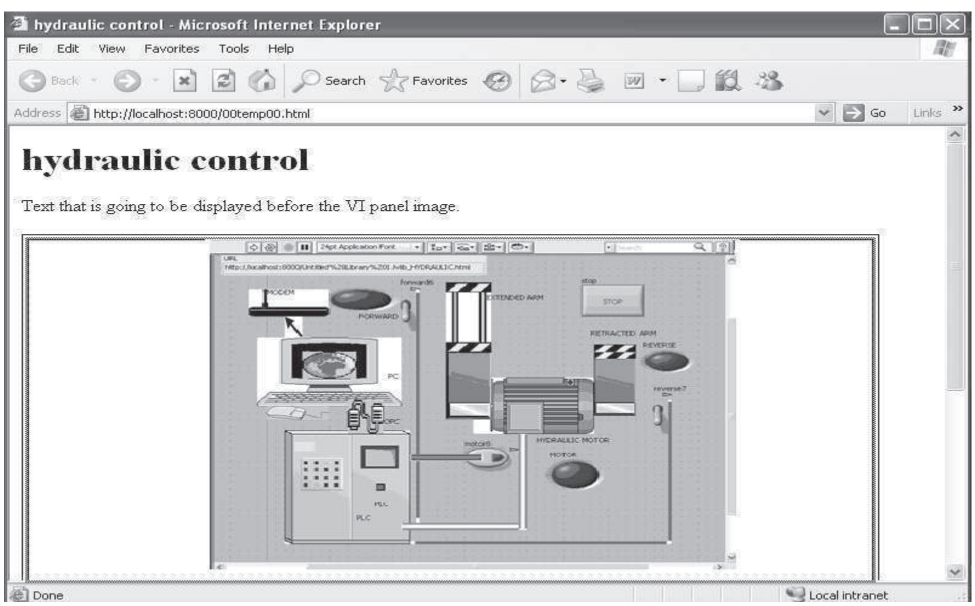


Figure 9. Remote monitoring and Control of Hydraulic Drive Control By Internet Access.

V. Implementation

Following steps are implemented to achieve desired objective to monitor and control Hydraulic drive/actuator.

1) Connecting PLC to the computer using an RS232 serial communication cable and also connecting sensor and actuator to PLC I/O module.
2) Configuring PLC Driver.
3) Creating new OPC Server
4) Creating a new Group
5) Creating Item/Tag
6) Creating new I/O variables.
7) Selecting data type.
8) Connect LabVIEW to OPC Tags by Creating an I/O Server.
9) Creating and Running VI as shown in Figure 7 and Figure 8.
10) Publishing this VI through web publishing tool.
11) Using created URL, monitoring and controlling the Hydraulic drive remotely. Thus now the control is transferred to remote place as shown in Figure 9.

Conclusion

Thus interfacing two most powerful technologies ruling the industries lead to many new features like acquiring data at faster rate, controlling the process accurately, remote controlling and several other features. Thus the PLC has been interfaced with LabVIEW through OPC Server and hydraulic drive system which is a real time processes have been controlled. Also the OPC enables a fully scalable solution for future changes and expansion. Thus the users will be no longer tied or locked in to a single vendor.

References

[1] Padhee. S and Singh.Y, “Data Logging and Supervisory Control of Process Using LabVIEW”, Students Technology Symposium, IEEE 2011.

[2] Jacob, M. “Real-time hardware-in-the-loop test platform for thermal power plant control systems”, Intelligent Systems and Informatics (SISY), 2011 IEEE 9th International Symposium on, pp. 495 – 500, 2011.

[3] S. Aleksandrov, S. Cajetinac, Z. Jovanovic, "Internet access to PLC with integrated web server", Proceedings of the 11th International Conference Research and Development in Mechanical Industry, RaDMI 2011, September, 15.-18., 2011, Sokobanja, Serbia, pp.425-430.

[4] Kai Zhou, Xudong Wang, Chao Zhang and Jian Liu, “Data Acquisition System Based on LabVIEW ABS Dynamic Simulation Test Stand”, International Conference on Information and Automation, IEEE 2010.

[5] Wenge Li, Zhenmei Li, Wei Liu, Peiyu Wei and Yayang Bur, “Remote Monitoring and Analysis System of Mine Safety Based on Virtual Instrument”, International Conference on TTest and Measurement, IEEE 2009.

[6] Li Pengfei, “Remote Control Laboratory Based On LabVIEW”, Intelligent Computation Technology and Automation, 2009. ICICTA '09. Second International Conference on, Volume: 4, pp.84-87.

Voice Guided Robot using LabVIEW

Harivardhagini S[1], Pranavanand S[2] and Ghali Bharadwaja Sharma[3]
[1] C.V.R College of Engineering, Ibrahimpatan, R.R District, Andhra Pradesh, India
Email:harivardhagini@gmail.com
[2]VNR Vignana Jyothi Institute of Engineering & Technology, Andhra Pradesh, India
Email:pranavanand_s@vnrvjiet.in
3 C.V.R College of Engineering, Ibrahimpatan, R.R District, Andhra Pradesh, India
Email:ghali.bharadwaj@gmail.com

***Abstract*—The main aim of this paper is to develop an efficient and low cost voice guided robot that navigates based on the voice signals given by the human users using LabVIEW 2011 for speech recognition and RF module for wireless transmission. The module is aimed at navigating in environments with noise and other disturbances very effectively. The speech recognition is based on LabVIEW 2011 .NET frameworks and the VI designed is highly flexible and can suit to large variety of applications where in the changes are to be in the microcontroller programming and the hardware requirements which are relevant to the application being designed. The communication to the hardware has been established via RS232 cables using microcontrollers to the transmission kit and the RF signal to communicate with the receiving kit mounted on the mobile robot. Unlike the costly DAQs these hardware communication are highly cost effective and the systems are flexible and can be altered to any utility. Adequate measures are taken to reduce the problems of external disturbances like noise etc. so that the module will understand the command even in very noisy environments. The purpose of these modules is to reduce the human loss due interventions in hazardous environments such as probation of life under debris, bomb detection in target sites etc.**

***Index Terms*—Speech recognition, LabVIEW 2011.NET frameworks, RF communication.**

I. Introduction

As the control action performed by understanding the voice commands of the humans brings in highly convenient interaction with robots, we have designed multi- utility voice recognition. The implementation of this voice guided robot involves both hardware and software, where in the general block diagram of the complete execution is as follows:

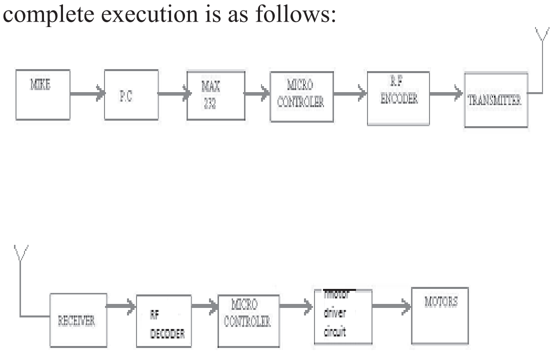


Figure 1.Block diagram

II. Description

The microphone will initially receive the voice signal from the human user. The output of this is received by the PC via 3.5mm jack. This received information is initiated onto to the LabVIEW using the appropriate constructor and is programmed to perform desired actions. The information from the PC has to communicate to the PCB or to the hardware for which a serial communication has to be established. This is done by using a USB to RS232 conversion cable, and the RS232 port is then connected to the MAX232 present on PCB which converts TT Logic to Binary Logic. The information from MAX232 is the received by the micro-controller and will transmit the signal according to program via RF Transmitter. The RF Receiver will receive the information and is given to the micro-controller and the necessary action taking by the controller by initiating the driver IC to drive the motors.

Thus as we give the voice commands the robot will move continuously by following the above algorithm.

III. Implementation

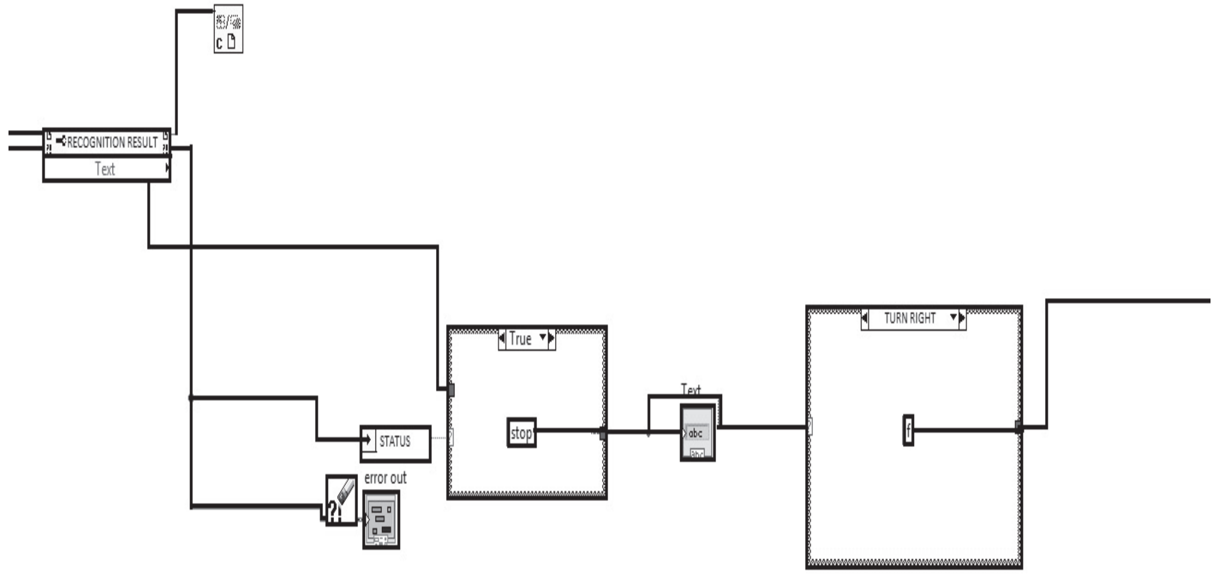
In this paper the LabVIEW programming and the micro controller coding are the major software programming involved. The LabVIEW programming involves the speech recognition and the speech to text conversion along with the serial communication.

A.The LabVIEW Programming

The constructor speech recognition engine is selected as the primary constructor and it is set up to a default audio device. Then it is initialized by the grammar to make the system understand the voice commands by the user.

The commands are predefined here to reduce the noise and other disturbances, the commands here are: Go Forward, Go Backward, Turn Left, Turn Right and Stop.

By the above commands the grammar is built. It is then connected to the text output. This text output is then given to case structure that will finally be displayed as the text box on the front panel of VI.Further this string output is connected to another case structure that will assign the action that is specified in the micro controller. The output of the above case structure is then connected to the serial communication part of the program that will write the information using the VISA serial communication modules present in LabVIEW. This is then given to the transmission kit that is connected to the PC which will finally transmit the encoded signal using the encoder.



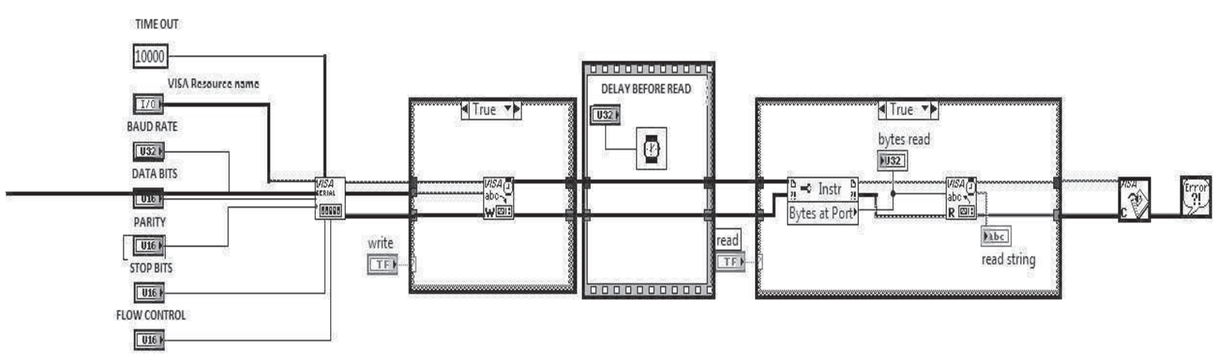


Figure 2. Block Diagram of module

IV. RF Module

The RF module was built with a transmitter and a receiver. Two micro controllers are used here, one in the transmitter section and the other in the receiving section. In the transmitter the micro controller is coded to receive the information the PC and transmit the same via RF Transmitter. The second microcontroller is coded to receive the information from the receiver and take the necessary actions using the driver IC to drive the motors.

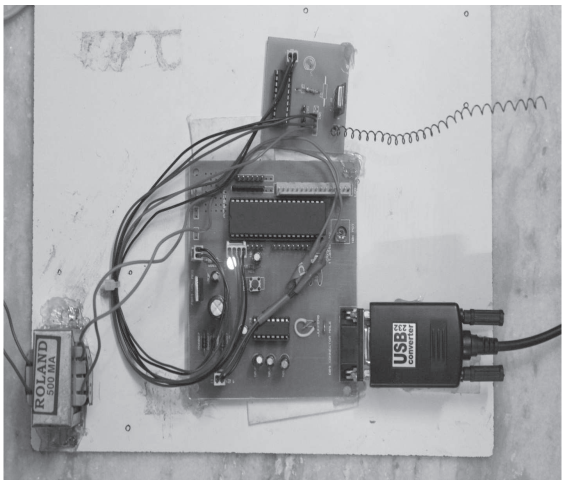


Figure 3. Transmitting side

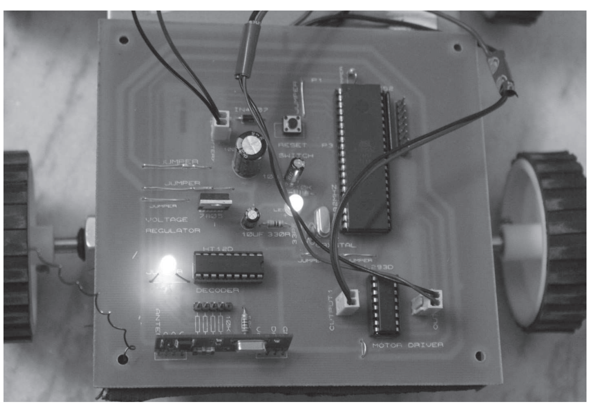


Figure 4 Receiving side

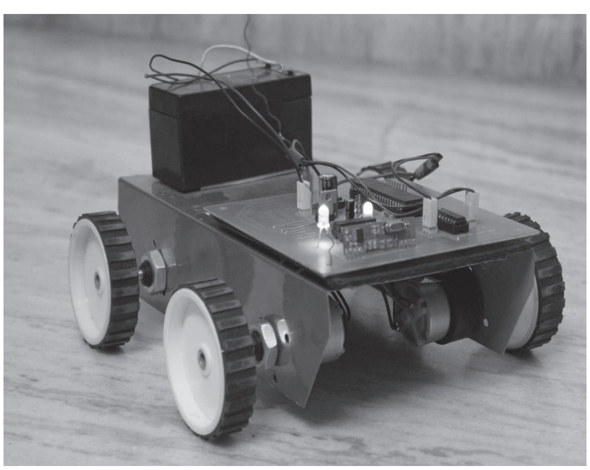


Figure 5. The Robot

This is the final prototype Robot with four metal gear dc motors connected to the wheels and the receiving kit is mounted over it. It has a battery (12V, 1.3Ah) mounted on the robot which the supply to both robot and the kit.

Result And Conclusion

The codes are deployed onto the micro controller and the serial communication is established. The voice commands are received and the robot navigates on the surface continuously for the commands specified.

The VI and the trans-receiving kit which is designed in this is aimed at lowest cost possible in the vicinage. This requires no specialized DAQs which are available in the market, and are very expensive too. This can be installed at various domestic, industrial, military, medical and many other fields as per the requirement.Though there are various hardware modules that performs this speech recognition, this LabVIEW based module is most advantageous as it is a software and can altered according then there innumerable times. Moreover the hardware will require higher maintenance and is prone to damage at hazardous situations. This module is highly flexible and can be used to perform any task by just making few changes in the commands and the hardware which has to be altered accordingly. Enhancements like addition of cameras, sensors, detectors etc. can be made suit the requirements. The programs can be run in lower platforms also, as well as on Desktops and Laptops. It is flexible on various OS environments like Microsoft windows, Apple Mac, Ubuntu (Linix) etc.

References

[1] Barry E. Paton 1999," LabVIEW Graphical Programming for Instrumentation". Prentice Hall PTP, New Jersey, U.S.A

[2] Curtis D. Johnson 1984," Microprocessor-Based Process Control". Prentice-Hall International, IncEnglewood Cliffs, NJ, U.S.A.

[3] DE-LORENZO, 2002, Electronic Laboratory,"(Basic Board To Study Temperature Regulation)".DL.2155RGT1, Dl.2155RGT2, Milano, Italy.

[4] Donald R. Coughanowr, 1991,"Process Systems Analysis and Control". McGraw-Hill, Inc., Singapore

[5] Gary W. Johnson, 1994. "LabVIEW Graphical Programming". McGraw-Hill, Inc., New York, U.S.A.

[6] J. Michael Jacob, 1989," Industrial Control Electronics". Prentice-Hall International, Inc., New Jersey, U.S.A.

[7] National Instruments, 2002, LabVIEW Basics Introduction, Course Manual, U.S.A

[8] National Instruments, 1996, "LabVIEW Graphical Programming for Instrumentation". User Manual. New Jersey, U.S.A.

Design and Development of Liquid Level Transmitter using an Improved Linearized Network

K.V.L.Narayana[1], Dr.A.Bhujanga Rao[2] and D.Linga Reddy[3]
[1]Gitam University, Visakhapatnam, A.P., India
Email: kvl007in@gmail.com
[2]Andhra University, Visakhapatnam, A.P., India
Email: dr_abrao@yahoo.co.in
[3]CVR College of Engineering, Department of H&S, Ibrahimpatan, R.R.District, A.P., India
Email: dlreddy_phy@yahoo.com

Abstract*—In this paper, a liquid level transmitter using cylindrical capacitive sensor and an improved linearized network for capacitance measurement has been proposed to measure the liquid level and to convert level changes into an electrical current which can be transmitted to a remote indicator. The change in capacitance of cylindrical capacitive sensor due to change in liquid level is measured by an improved linearized capacitance measuring network. The offset capacitance of the cylindrical capacitive sensor and the stray capacitances that exist between sensor electrodes & metallic tank are measured before the liquid level measurement. The measured capacitances are used in the proposed capacitance measuring network to minimize the effects of offset capacitance and stray capacitances on liquid level measurement using dc control voltage and operational amplifiers with high input impedance. The experimental investigations have been performed to sense water level of metallic tank in both increased and decreased level conditions. In the first phase of experiment, a linearized network has been simulated using LabVIEW (laboratory virtual instrument engineering workbench) and studied with the test capacitance, and in the second phase, the experimentation was done by replacing the test capacitance with a cylindrical capacitive sensor for the measurement of liquid level. As a result of investigations conducted, it has been observed that the variation in liquid level from 0 to 25cm having linear relationship with output dc voltage in the range of 0 to 5.5V. Corresponding to liquid level variations, the dc output voltage further converted into an electric current of 4 to 20mA for remote indication and control purpose. The experimental results of liquid level transmitter are found to have good linearity of about ± 0.2% and a resolution of about 1 cm. The sensitivities of the capacitance measuring circuit and level transmitter have been found about 6.5 mV/pF and 250mV/cm respectively.

***Index Terms*—Cylindrical capacitor, LabVIEW, Capacitance measurement, Phase sensitive detector, Linearization.**

I. Introduction

Sensor technology is in the process of a slow migration from discrete dumb instruments, expensive and inflexible, to smart, self-calibrating, silicon-based units, and the measurement method of choice for discrete instruments is moving from a variety of transducer technologies, such as magnetic, optical, and piezoelectric, to capacitive. Capacitive sensors [1] electronically measure the capacitance between two or more conductors in a dielectric environment, usually air or a liquid. These sensors can directly sense a variety of things—motion, chemical composition, electric field—and, indirectly, sense many other variables which can be converted into motion or dielectric constant, such as pressure, acceleration, fluid level, and fluid composition.

Cylindrical capacitive sensor had been originally introduced by Chapman [2] for its advantages, which are the insensitivity to geometric errors by the averaging effect, which is derived from a simple intuition that the summation of geometric errors eventually converges to zero since the mean of geometric errors is zero and the high resolution with large sensing area compared to probe-type sensors.

In any process industry, liquid such as water, chemicals, and solvents in a storage vessel is required to be measured and controlled. The amount of such liquid stored can be found by measuring liquid level in a container or vessel. The liquid level affects the quantity delivered in and out of the container. The pressure and flow rate of liquid affects the liquid level of the container.

The change in capacitance of a capacitive sensor due to a change in process variable is generally very small. Hence, various attempts have been made by different investigators [3]–[17] to accurately measure this change in capacitance. In the conventional bridge methods, the Schering bridge [11] is best suited for the measurement of capacitance but the bridge methods are tedious and time consuming, as convergence toward balance requires several iterative steps. Automatic balancing bridges [12]–[14] have been developed but with increased complexity and cost. The direct-reading technique of capacitance measurement reported in [15] though useful, requires involved computation for determining the parameters from measured voltages. The microprocessor based switched-battery capacitance meter proposed in [16] is simple, but the measured values do not reflect the accepted equivalent circuit parameters applicable to sinusoidal excitation. In the method reported [17], the unknown capacitance changes the frequency of an oscillator and the frequency deviation is used as a measure of the capacitance. A new method of measuring capacitance, using oscillator circuits [18] requires a standard capacitor of value nearly equal to that of the nominal value of the unknown capacitor. The above two methods have a limitation that the unknown capacitor cannot be tested at a desired frequency and voltage. A simple scheme for the measurement of capacitance and dissipation factor of capacitor proposed in [19] has less resolution and hence it is not suitable for low value capacitance measurement. In this scheme, the accuracy of the prototype unit for the measurement of capacitance has been found as ±0.2%. Hence, it is aimed to improve the linearity of capacitance measuring circuit of about ±0.1% and to compensate the stray and parasitic capacitances exist in the capacitive sensor, capacitance measuring circuit.

This paper proposes a design of virtual instrument system for the measurement of small level changes using improved linearized network for capacitance measurement. A uniform right circular cylinder made of polyvinyl chloride has been used as a cylindrical capacitive sensor in the designed virtual instrument system. The offset capacitance of the cylindrical capacitive sensor and the stray capacitances that exist between sensor electrodes & metallic tank are measured before the liquid level measurement. The measured capacitances are used in the proposed capacitance measuring network to minimize the effects of offset capacitance and stray capacitances on liquid level measurement using dc control voltage and operational amplifiers with high input impedance. In this instrument, resolution and linearity have been improved. The output current has been found to be linearly related to the level changes. The capacitance changes have been measured and interfaced to PC-LABVIEW through NI-PXI-4072. The proposed system is easy to implement and convenient for various applications.

II. Principle of Cylindrical Capacitive Sensor

Consider a cross section of solid cylindrical conductor of radius r_1 surrounded by a coaxial cylindrical shell of inner radius r_2. The length of both cylinders is L which is to be much larger than $r_2 - r_1$, the separation of the cylinders, so that edge effects can be neglected. The capacitor is charged so that the inner cylinder has charge $+Q$ while the outer shell has a charge $-Q$. To calculate the capacitance, the electric field produced by charge on cylinder is computed. Due to the cylindrical symmetry of the system, gaussian surface is to be chosen as a coaxial cylinder with length $l<L$ and radius r where $r_1<r<r_2$.

Using Gauss's law,

$$\oiint \vec{E}\cdot d\vec{A} = EA = \frac{\lambda l}{\varepsilon_0} \qquad (1)$$

$$E = \frac{\lambda}{2\pi\varepsilon_0 r} \qquad (2)$$

Where $\lambda = Q/L$ is the charge per unit length; ε_0 is the absolute permittivity of free space. It is to observed that the electric field is non-vanishing only in the region $r_1<r<r_2$. For $r< r_1$, the enclosed charge is zero since any net charge in a conductor must reside on its surface. Similarly, for $r>r_2$, the enclosed charge is $\lambda l - \lambda l=0$, since the gaussian surface encloses equal but opposite charges from both conductors.

Therefore, the potential difference between two cylindrical electrodes is given by

$$\Delta V = V_{r_2} - V_{r_1} = -\int_{r_1}^{r_2} E_r dr = -\frac{\lambda}{2\pi\varepsilon_0}\ln\left(\frac{r_2}{r_1}\right) \qquad (3)$$

The change in capacitance, ΔC with respect to the liquid column of height, h_1 and remaining height, h_2 of the metallic tank with air as the dielectric is given by

$$\Delta C = \frac{2\pi\varepsilon_0(\varepsilon_1 h_1 + \varepsilon_2(L-h_1))}{\ln\left(\frac{r_2}{r_1}\right)} \qquad (4)$$

Where h_1 is the height of liquid column, L is length of the cylindrical capacitor, r_1 is the radius of solid cylindrical conductor , r_2 is the inner radius of coaxial cylindrical shell and ε_0, ε_1 & ε_2 are the permittivity of free space, liquid and air respectively.

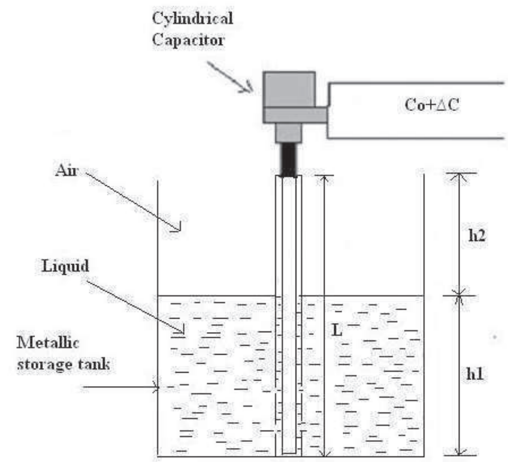


Figure 1. Liquid filled metallic storage tank with cylindrical capacitor.

The cylindrical capacitor immersed in the liquid filled metallic storage tank and it touches the metallic storage tank as shown in figure 1. The inner conductive cylinder of the capacitive sensor is filled with the air. There is another capacitance C_0 between the metallic storage tank or other metallic objects and the cylindrical capacitor. This capacitance may be assumed to be the parallel combination of parasitic capacitances exists: between metallic vessel and upper part of the probe above the liquid level; between metallic vessel and lower part of the probe below the liquid level. This capacitance in the order of pF may be assumed to be connected in parallel with the test capacitance between the liquid column of height h_1 and the sensing probe. The stray capacitance which exists between the electrodes and measuring circuit is independent of liquid level to be measured. It alters the effective values of measuring circuit components. The effective capacitance of the sensor consists of the parasitic and stray capacitances. Hence, the effective capacitance of the sensing probe, C_s with respect to the liquid column of height, h_1 may be given by

$$C_s = C_0 + \Delta C \qquad (5)$$

Combining (4) and (5), we have

$$C_s = C_0 + \frac{2\pi\varepsilon_0(\varepsilon_1 h_1 + \varepsilon_2(L - h_1))}{\ln(\frac{r_2}{r_1})} \qquad (6)$$

Or $$C_s = k_1 h_1 + k_2 + C_0 \qquad (7)$$

Where $k_1 = \frac{2\pi\varepsilon_0(\varepsilon_1 - \varepsilon_2)}{\ln(\frac{r_2}{r_1})}$ and

$$k_2 = \frac{2\pi\varepsilon_0\varepsilon_2 L}{\ln(\frac{r_2}{r_1})} \qquad (8)$$

Thus, (7) indicates that C_s is linearly related to the liquid level, h_1.

III. Measuring Circuit

An improved linearized network using op-amps and phase sensitive detector has been designed as shown in figure 2 to measure the capacitance of the cylindrical capacitive sensor. The amplifiers, A_1 and A_2 are two similar TLC 274CN Precision quad operational amplifiers. The phase sensitive detector is a high precision balanced modulator/demodulator, AD-630 that combines a flexible commutating architecture with the accuracy and temperature stability afforded by laser wafer trimmed thin film resistors and the multiplier is a wideband linear four-quadrant multiplier, MC-1495. A sinusoidal voltage V_{in}, is applied to the capacitance C_s to be measured and also it is applied to the multiplier. A variable dc control voltage, V_c which is a function of V_{in} is connected to the multiplier as other input. For an excitation voltage, V_{in} the current passing through C_s and C_m would be i_s and i_m respectively.

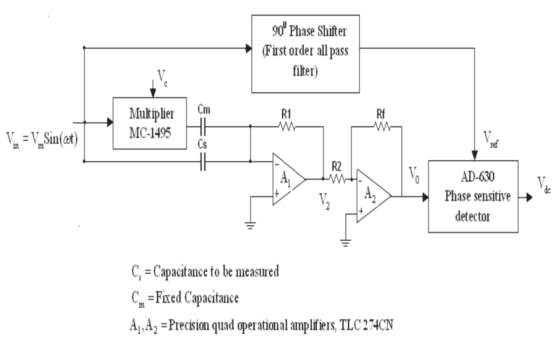


Figure 2. An improved linearized network for capacitance measurement.

The output voltage, V_{dc} of the circuit is a linear function of the capacitance, C_s as established by the following expressions.

$$V_{in} = V_m \text{Sin}(\omega t) \qquad (9)$$

$$V_2 = -R_1 C_s \frac{dV_{in}}{dt} - R_1 C_m \frac{dV_c}{dt} \qquad (10)$$

Let the control voltage V_c be the function of input voltage V_{in}

$$V_c = kV_{in} \qquad (11)$$

Where k is the constant and selection of its value depends on the offset capacitance of the sensor.

Therefore using equations (9)-(11), the outputs of A_1 and A_2 Op-amps, V_2 and V_O respectively can be written as,

$$V_2 = -R_1\omega V_m(C_s + kC_m)\text{Sin}(\omega t + 90^0) \qquad (12)$$

$$V_O = \frac{R_f}{R_2}(C_s + kC_m)R_1\omega V_m \text{Sin}(\omega t + 90^0) \qquad (13)$$

V_{in} is phase shifted by 90^0 and is used as the reference signal V_{ref} for the phase sensitive detector. The output of phase sensitive detector which includes the rectification, amplification and filtering is given as

$$V_{dc} = \frac{4R_f}{\pi R_2}(C_s + kC_m)R_1\omega V_m \qquad (14)$$

For a fixed value of capacitance C_m, excitation frequency ω, R_1, R_2, R_f, the output of phase sensitive detector is a dc voltage directly proportional to C_s to be measured. The effect of stray capacitances between the capacitor electrodes and the ground on the measurements is very small due to use of high input impedance op-amps in the measuring circuitry. The effects if any can be subtracted in the measurement by adjusting the control voltage, V_c.

The equations (7) and (14) give

$$V_{dc} = \frac{4R_f R_1 \omega V_m}{\pi R_2}(k_1 h_1 + k_2 + C_0 + kC_m) \qquad (15)$$

Where, the value of k is used to compensate the offset capacitance, C_0 of the sensor. Hence, C_0 and kC_m are cancelled each other. If the V_{dc} is the circuit output voltage, then

$$V_{dc} = \alpha((\varepsilon_1 - \varepsilon_2)h_1 + \varepsilon_2 L) \qquad (16)$$

Where $$\alpha = \frac{8\varepsilon_0 R_f R_1 \omega V_m}{R_2 \ln(\frac{r_2}{r_1})} = \text{constant} \qquad (17)$$

That is the output voltage is linearly related to change in the liquid level if the capacitive transducer is linear. The equation (16) depicts the linear relationship between the output voltage and changes in level, h_1. The output voltage signal is amplified to a voltage signal, V^1_{dc} in the range of 1–5 V dc, which is finally converted into a current signal (I_0) in the range of 4–20 mA dc by a voltage to current converter. After calibration, the output of the level transmitter becomes 4 mA when V^1_{dc} is 1 V and liquid level h_1 is zero cm and becomes 20 mA when V^1_{dc} is 5 V and liquid level h_1 is at the maximum range (h_{1max}). Hence, the level transmitter voltage output, V^1_{dc} in volts and the current output (I_0) in milliamperes may be written as

$$V^1_{dc} = \left(\frac{4}{h_{1\max}}\right)h_1 + 1 \qquad (18)$$

$$I_0 = \beta V^1_{dc} \qquad (19)$$

Where h_{1max} is the maximum value of liquid level selected as 25cm and β is constant. From equations (18) and(19)

$$I_0 = \beta + \left(\frac{4\beta}{h_{1\max}}\right) h_1 \tag{20}$$

An improved linearized network using op-amps and phase sensitive detector has been used for measuring capacitance, replacing the variable capacitor by a cylindrical capacitor immersed in metallic tank. The metallic tank was fitted with a graduated scale so that level can be measured.

Since from (16), $h_1 \cong (\frac{1}{\varepsilon_1 - \varepsilon_2})\frac{V_{dc}}{\alpha}$ for small value of $\varepsilon_2 L$, the relative measurement error E_1 expressed in percentage may be calculated by the following equation:

$$E_1 = \frac{((\frac{1}{\varepsilon_1 - \varepsilon_2})\frac{V_{dc}}{\alpha} - h_1)}{h_1} \times 100\% \tag{21}$$

Where h_1 is the measured level; V_{dc} is the measured output voltage.

In terms of percentage of the maximum range, it changes to (E_2); this may be defined as relative measurement error

$$E_2 = \frac{((\frac{1}{\varepsilon_1 - \varepsilon_2})\frac{V_{dc}}{\alpha} - h_1)}{h_{1\max}} \times 100\% \tag{22}$$

Where h_{1max} is the maximum value of liquid level change. Again, from the measured values of V_{dc} at different values of h_1, the best-fit linear characteristic may be drawn using the LabVIEW. The linear least squares fitting technique which is the simplest and most commonly applied form of linear regression, has been used to find the best-fitting curve to a given set of points (measured values of V_{dc} at different values of h_1 by minimizing the sum of the squares of the residuals of the points from the curve. The actual values of V_{dc} be obtained from the best-fit linear characteristic at different values of h_1, and the relative output voltage measurement error E_3 from linearity may then be defined as:

$$E_3 = \frac{(V_{dc-actual} - V_{dc})}{V_{dc-actual}} \times 100\% \tag{23}$$

Where $V_{dc\text{-}actual}$ is the actual values obtained from the best-fit linear curve for a given level change, h_1.

The virtual instrument for the measurement of small liquid level changes is shown in figure 3. The output voltage of capacitance measuring circuit has been filtered and interfaced to PC-LabVIEW through NI-cDAQ 9174 in order to transmit over remote location devices. The virtual instrument provides low-cost and flexible solution for the acquisition of signals as well as transmission of analyzed signals over remote location.

IV. Experimental Results

The experiment was performed in two phases. In the first phase of the experiment, an improved linearized network for capacitance measurement performance has been studied with a known variable capacitor and ac test signal of sinusoidal voltage ($V_{in}=V_m \sin(\omega t)$), with an amplitude V_m of 5V and a frequency of 1.0 kHz. The values of components used in the measuring circuit (as shown in Fig.2.) are R_1=100kΩ, R_2=10kΩ, R_f=15kΩ, C_m=22pF. The amplifiers, A_1 and A_2 are TLC274CN, precision quad operational amplifiers, the phase detector is balanced modulator/demodulator, AD-630 and the multiplier is a wideband linear four-quadrant multiplier, MC-1495. The 4 (4/5) digit TX3 true digital multimeter was used for measuring output voltage. An ac waveform of a particular frequency and a 90 degrees phase shifted waveform of the same frequency are applied to the signal and reference inputs of the phase sensitive detector, AD-630. The dc level of the phase detector output is proportional to the signal amplitude and phase difference between the input signals. If the signal amplitude is held constant, the output can be used as a direct indication of the phase. When these input signals are 90° out of phase, they are said to be in quadrature and the dc output of AD-630 will be zero.

The change in capacitance, ΔC in both increasing and decreasing modes has been varied in steps of 1 pF over 10pF range and at each step, the output voltage was measured. The experiment was repeated for different values of control voltage, V_c based on values of stray capacitance as shown in table 1. Three values of the stray capacitances are assumed in table 1 and the corresponding control voltages are obtained from the measuring network by considering the both sensor offset capacitance and the assumed stray capacitances. The control voltages at the input of multiplier are measured in such a way that which makes the measuring network output zero in absence of the test capacitance. Experimental characteristic graphs were then drawn by plotting the output voltage against the known variable capacitance, ΔC for different values of the control voltage, V_c as shown in figure 4. The error bar was drawn for the measurement points of output voltage and change in capacitance as shown in figure 5. From these experimental data the best fit straight-line curve was plotted by using LabVIEW 9 in each case, and the percentage error of the experimental data from this optimum straight line was calculated. The percentage deviation curve of the change in capacitance from a straight line for different values of V_c is drawn by using equation (23), as shown in figure 6.

In the second phase of experiment, the improved linearized circuit for capacitance measurement has been used for measuring level, replacing the variable capacitor by a cylindrical capacitor immersed in liquid filled metallic tank. The cylindrical capacitor of the virtual instrument system is the laboratory standard equipment with the following specification.

L: Length of the cylinder, 300mm

r_1: Internal radius, 4mm

r_2: External radius, 15mm

ε_1: Dielectric constant of water, 80.4 (at 20^0C)
ε_2: Dielectric constant of air, 1.0548 (at 20^0C
ε_0: Permittivity of free space

A. Material: High density polyethylene

The liquid filled metallic tank was fitted with a graduated scale in cm so that level can be measured. The offset capacitance of the cylindrical capacitive sensor with zero level is C_{in} = 12.6pF and it has been assumed that parasitic capacitances in the measurement as C_0=10pF. Hence, the base capacitance of capacitive sensor with zero level is approximately 22pF. Initially, the control voltage of the measuring circuit is adjusted to get zero output voltage and then, the tank is filled by a tap water having a resolution of 1cm. The output voltages for different values of the level with a selected value of V_c which depends on the value of stray capacitance, are measured in both increasing and decreasing modes. The variation of the output voltage with the change in liquid level is found to be linear as shown in figure 7. The static characteristic graph of the level transmitter was drawn by plotting transmitter output current of 4 to 20mA against liquid level variation from 0 to 25cm. as shown in figure 8. The linear characteristics over a wide range of level with good linearity, and resolution have been described.

Table I.
Control Voltage For Various Values Of Stray Capacitances.

S.No.	Sensor offset capacitance (pF)	Stray capacitance (pF)	Control Voltage, V_c (V)
1.	12.6	10	2
2.	12.6	5	1
3.	12.6	0	0

CONCLUSIONS

The basic materials for developing this liquid level transmitter are simple high density polyethylene cylindrical capacitor and the signal conditioning circuit which involves low-cost semiconductor devices. The construction technique does not involve any high-cost technology compared with existing non-contact type level sensors such as ultrasonic gauge and nuclear absorption level gauge. From the experimental study, the repeatability, linearity, and resolution are satisfactory within the tolerable limit of industrial level measurement. Hence, the present technique may be treated as a low-cost linear alternative technique of level measurement of both conducting and non conducting liquids.

The experimental characteristics of the proposed level transmitter shown in figures 4 and 7 for capacitance and level measurement, respectively, have been found to be quite linear about ±0.1% within a resolution of about 1 cm. The sensitivities of measuring circuit and level transmitter have been found about 6.5mV/pF and 250mV/cm respectively. The percentage deviation of capacitance measurement from the best fit linear graph, as shown in figure 6 has found to be within the tolerable limit. The human error in taking the reading of the level may contribute to a small percentage error as the level has been measured by a graduated scale. The offset capacitance of the cylindrical capacitive sensor and the stray capacitances that exist between sensor electrodes & metallic tank are measured and have been used in the proposed capacitance measuring network to minimize the effects of offset capacitance and stray capacitances on liquid level measurement using dc control voltage and operational amplifiers with high input impedance. The static & dynamic uncertainty analysis of level sensing technique could not be performed. From this uncertainty analysis, the proposed improved signal conditioning technique may be further modified in order to reduce the uncertainty of the liquid level measurement. The proposed virtual instrument for level measurement has very good accuracy and resolution compared to the conventional measurements available in literature.

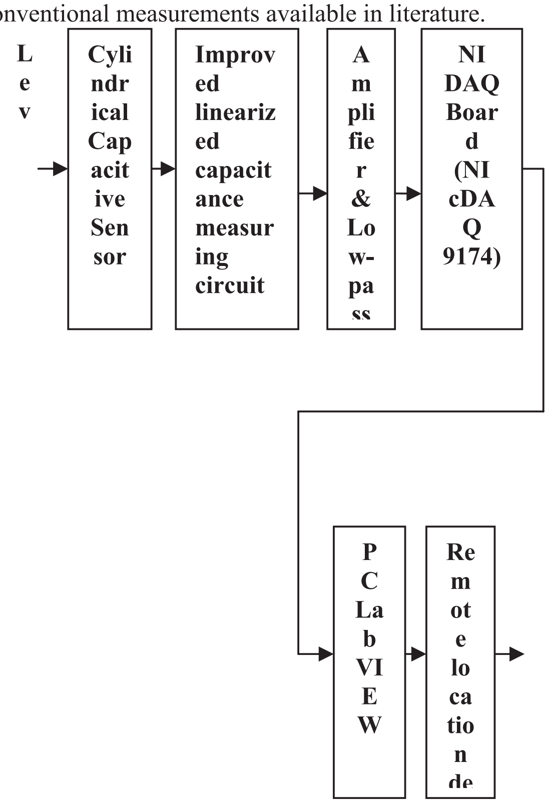


Figure 3. Block diagram of virtual instrument system for level measurement.

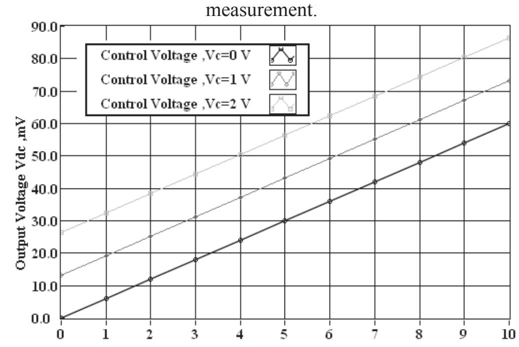


Figure 4. Change in linearized circuit output voltage versus variation in capacitance.

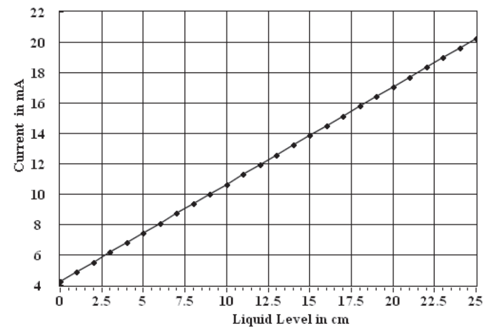


Figure 5. Error bar for measurement points of output voltage and change in capacitance.

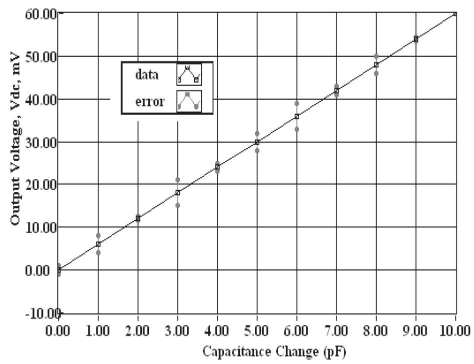


Figure 6. Percentage deviation of the capacitance from straight line characteristic.

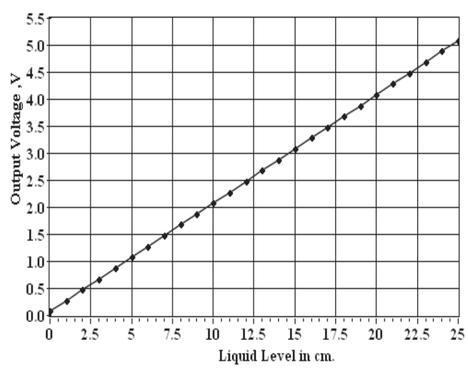


Figure 7. Variation in linearized circuit output voltage with change in liquid level.

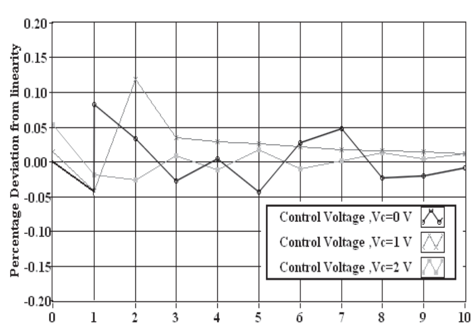


Figure 8. Variation in output current with change in liquid level of level transmitter.

REFERENCES

[1] L. K. Baxter, Capacitive Sensors, Design and Applications, IEEE Press, New York (1997).

[2] P. D. Chapman, *A capacitive based ultra precision spindle error analyzer*, *J. Precision Eng.* 7 (1985) 137.

[3] B. G. Liptak, *Process Measurement and Analysis*, Oxford, U.K. (1999).

[4] D. M. Considine, *Process Instruments and Control Hand Book*, McGraw-Hill, New York (1974).

[5] A. J. Jaworski, T. Dyakowski, and G. A. Davies, *A capacitance probe for interface detection in oil and gas extraction plant*, *Meas. Sci. Technol.* 10 (1999) L20.

[6] W. R. Yang, M. R. Brant, and M. S. Beck, *A multi-interface level measurement system using a segmental capacitance sensor for oil separators*, *Meas. Sci. Technol*. 5 (1994)1180.

[7] S. M. Herang, R. G. Green, A. Plaskowski, and M. S. Beck, *A high frequency stray immune capacitance transducer based on the charge transfer principle*, *IEEE Trans. Instrum. Meas*.37 (1998) 373.

[8] A. H. Falkner, *The use of capacitance in the measurement of angular and linear displacement*, *IEEE Trans. Instrum. Meas*.43 (1994) 942.

[9] P. Holmberg, *Automatic balancing of linear AC bridge circuits for capacitive sensor elements*, *IEEE Trans. Instrum. Meas.* 44 (1995) 805.

[10] D. Marioli, E. Sardini, and A. Taroni, *High accuracy measurement techniques for capacitance transducers*, *Meas. Sci. Technol.* 4 (1993) 343.

[11] Foley A.H, A direct reading high-voltage capacitance bridge, Transactions of American institute of electrical engineers.69 (1950) 698.

[12] O. Petersons, *A self-balancing high voltage capacitance bridge*, *IEEE Trans. Instrum. Mea.*13 (1964) 224.

[13] O. Petersons and W. E. Anderson, *A wide-range high-voltage capacitance bridge with one ppm accuracy*, *IEEE Trans. Instrum. Meas.* 24 (1975) 344.

[14] R. D. Cutkosky, *An automatic high-precision audio frequency capacitance bridge*, *IEEE Trans. Instrum. Meas*.34 (1985) 389.

[15] M. S. Ansari and M. T. Ahmed, *A novel direct reading active-RC system for measurements of in-circuit, discrete, and incremental capacitances*, *IEEE Tran. Instrum. Meas*.38 (1989) 992.

[16] S. M. Mahmud and A. Rusek, *A microprocessor-based switched battery capacitance meter*, *IEEE Trans. Instrum. Meas*.37 (1988) 194.

[17] T. Takagi and A. Yamakawa, *A simple and wide-range capacitance measuring equipment using transistor blocking oscillator*, *IEEE Trans. Instrum. Meas.* 25 (1976) 163.

[18] S. Natarajan, *Measurement of capacitance and their loss factors*, *IEEE Tran. Instrum. Meas*.38 (1989) 1087.

[19] V. Jagadeesh Kumar, P. Sankaran, and K. Sudhakar Rao, *Measurement of C and tanδ of a capacitor employing PSDs and Dual-slope DVMs*, *IEEE Trans. Instrum. Meas*.52 (2003)1592.

Investigation of Dielectric Properties of (1-x) $Ni_{0.53}Cu_{0.12}Zn_{0.35}Fe_{1.88}O_4$ ferrite + (x) $Gd_{0.2}Ce_{0.8}O_3$ for Multilayer Chip Inductors

S.Bharadwaj[1] ,Suman Kumar Burnwal[2], T.Ramesh[3] and S.R.Murthy[4]
[1,2] CVR College of Engineering, Department of H&S, Ibrahimpatan, R.R.District, A.P., India
[3,4] Osmania University, Hyderabad, India, A.P., India

Abstract: **The composite of (1-x) $Ni_{0.53}Cu_{0.12}Zn_{0.35}Fe_{1.88}O_4$ + (x) $Gd_{0.2}Ce_{0.8}O_3$(x=0.10,0.20,0.30) were prepared by mixing nanocrystalline $Ni_{0.53}Cu_{0.12}Zn_{0.35}Fe_{1.88}O_4$ and $Gd_{0.2}Ce_{0.8}O_3$powders at different weights percents. The powders of NiCuZn ferrite were synthesized using sol gel method. The powders were densified using conventional sintering method at 1000°C/2 hrs. The phase and morphology of the composites was observed with X-ray diffraction (XRD) and Scanning Electron Microscope (SEM). The frequency dependence of real (ε′) and imaginary (ε″) parts of permittivity was measured in the range of 1MHz-1.8GHz.**

Index Terms**—Composites, X-ray diffraction, Electrical properties.**

I. INTRODUCTION

Ferrites are usually as core materials in inductor which behaves as an impedance device and has a coil for inducing inductance to an electronic circuit. A ferrite core typically amplifies or multiplies the inductance of a coil by the permeability of the core. The chip inductor is fabricated by laminating ferrite layers and internal conductors and co-firing for monolithic structure [1]. Ferrite chip inductors are one of the important components for majority of electronic products such as cellular phones, computer notebooks, hard and floppy drives etc [2]. The decrease of sintering temperature along with increase of densification of the material are desirable conditions such that it can be co-fired with silver (Ag) internal electrode for MLCI application. The various concentration of zinc and copper and optimization of zinc are essential to achieve desirable electromagnetic properties in the ferrites those have been reported in the literature [3]. This makes Ni substituted Zinc ferrites dominant materials for multilayer chip inductor (MLCI) applications as they have good electromagnetic properties. With development of communication technology, the high performance MLCIs used in the frequency range from 500 MHz to 2 GHz. The MLCIs should have good density, high permeability and low dielectric constant assuring the high resonance frequency. The addition of ferroelectric material such as $BaTiO_3$ with NiCuZn ferrites resulted increase in the cutoff frequency [4]. The rare earth substituted materials are becoming the promising materials for different applications. Addition of small amount of rare earth ions to ferrite samples produces a change in electrical and structural properties depending on the type of substitution. Zhao et al. [5] investigated that the crystallite size of the Ni-Mn ferrites decreased as Gd ions were doped. Ahmed et al. [6] reported the electrical properties of Mg-Ti doped ferrites like doped with rare earth ions Er, Ce and Nd increased the resistivity. Sun et al. [7-8] investigated the effects of rare earth ions on the properties of $(Ni_{0.5}Zn_{0.5})Fe_{1.98}RE_{0.02}O_4$ (RE = Y, Eu or Gd) nominal compositions. The partial substitution of rare earth materials increased the density, higher electrical resistivity and relative loss factor. Hence in the present study, the NiCuZn ferrite particles were mixed with GdC in different weight percentages and sintered to form composite ferrite consisting of ferrite and rare earth materials. This paper reports the preparation, characterization and dielectric properties of rare earth doped ferrites composites.

II. EXPERIMENTAL

Analytical grade nickel nitrate, zinc nitrate, copper nitrate, iron nitrate and citric acid were used as raw materials to synthezied the $Ni_{0.53}Cu_{0.12}Zn_{0.35}Fe_{1.88}O_4$ ferrite material. The stoichometric proportions of molar ratio 1:1 of metal nitrates and citric acid is dissolved in the deionized water. During this process, the solution was continuously stirred using a magnetic agitator. Ammonia was added to the solution to adjust the pH value to about 9. The precipitate is then heated continuously until all the gel burnt out to form powder. The powder is then crushed using agate mortor and then heated at 65^{o}C for 12 hours. The particle size was estimated to be around 35 nm for NiCuZn ferrite using XRD. The $Gd_{0.2}Ce_{0.8}O_3$ (GdC)is directly purchased from the Cottor international manufacture with specified grain size of 50 nm and 99.9% purity.

The ferrite powder and (GdC) material is mixed in different weight percentages of x=0.10, 0.20, 0.30. The powder was uniaxially pressed at a pressure of 1500 kg/cm^2 to form green pellets. The pellets were sintered at 1000^{o}C and characterization is done using X-ray diffraction (XRD) and Scanning electron microscope (SEM). The phase identification of the sintered samples was performed using X-ray diffraction (XRD) with Cu-Ka radiation. The frequency dependence of complex permittivity was measured in the range of 1MHz to 1.8GHz using Agilent 4291B Impedance/Material Analyzer.

III. RESULTS AND DISCUSSION

X-ray diffraction patterns of composite ferrite of NiCuZn ferrite and (GdC) materials are shown in Figure 1., indicating that the sintered powders show the peaks for both the materials. Fig.1 also indicates the presence of two crystallite phases of NiCuZn ferrite and both (GdC) which coexist in all of three sintered composites. The lattice constants of composite material is estimated and tabulated in Table 1. from the table, it can be observed that the lattice constants decreases with increase in rare earth dopants. It was also found that the density of the present composites vary 90% that of theoretical density and hence the values of porosity vary around 10%.

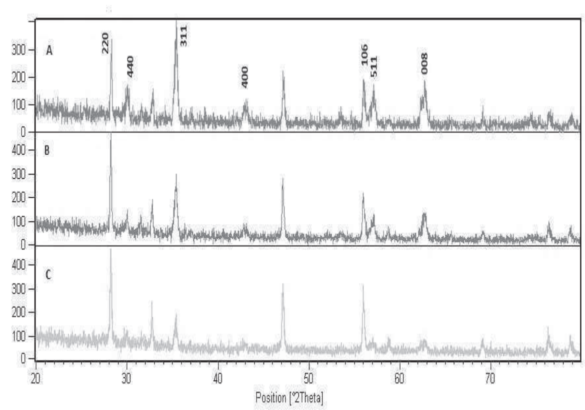


Figure 1. XRD of the composites

Sample Name	Lattice constant		Grain Size (SEM) nm	ε′ (1GHz)	ε″ (1GHz)
	a(A°) ±0.001	c(A°) ±0.001			
A	5.351	7.656	92	19.64	0.384
B	5.276	7.456	104	19.24	0.337
C	5.125	7.223	117	18.00	0.318

Table I. Data of lattice constant, grain size and electrical properties

Figure 2. shows the SEM pictures for conventional sintered samples. The average size of the composite ferrite, geometrically estimated from SEM photographs of composite sintered samples and results are presented in Table I. In the images, the black grains are ferrite grains and the white ones are ferroelectric grains. With increasing ferrite content, the ferrite grains increase and the ferroelectric grains decrease continually.

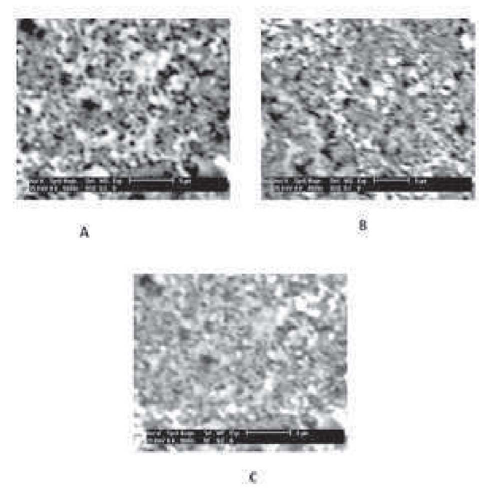


Figure 2. SEM for composites

The decrease of lattice constant and increase in grain size as the rare earth substance increases is due to the deformation of the structure. The difference in their ionic radii will lead to strains, which may result in domain wall motion resulting in deformation of the spinel structure [9]. The rare earth ions tendency to prefer the octahedral sites by replacing iron ions (Fe^{3+}) [10]. Ferrimagnetism is largely governed by Fe-Fe interaction (the spin coupling of the 3d electrons). If the rare earth ions enter the spinel lattice, the R_E-Fe interactions also appears (4f-3d coupling), which can lead to changes in the structural, electrical, magnetization and Curie temperature [11].

The frequency variation of real (ε′) and imaginary (ε″) parts of permittivity for all the samples under investigation was measured in the frequency range of 1 MHz to 1.8 GHz and obtained results are plotted in Fig.3 & Fig.4. It can be seen from the figures that the value of ε′ remains constant upto a frequency 600 MHz and increases further increase of frequency. In all the samples a resonance peak and anti resonance was observed above 1.49 GHz. This behaviour can be explained in the following way: The ε' remains constant in the frequency range from 1 MHz to 600 MHz due to the hoping electrons will not follow the external applied field. Where as the increase of ε' from 600 MHz to 1.49 GHz is may be due to the following of hopping electrons, the external field. When the hoping frequency of the electrons is equal to that of the external applied electric field, a dielectric resonance peak is obtained in the dielectric constant. The decrease of real part of permittivity (ε′) is attributed to the introduction of $Gd_{0.2}Ce_{0.8}O_3$which decreases the dipoles in the composite materials.

Frequency dispersion of permittivity for composites decreases with increasing rare earth dopants content due to the decrease in relaxor behavior caused by the incorporation of the ferrites [12]. The decrease in permittivity and broadness observed is due to incorporation of rare earth dopants which dilutes ferroelectric properties which can be observed from the figures. It can be observed from the figure that as the rare earth dopant increases in the composite materials, the ferroelectric transition is shifting towards the higher frequency side.

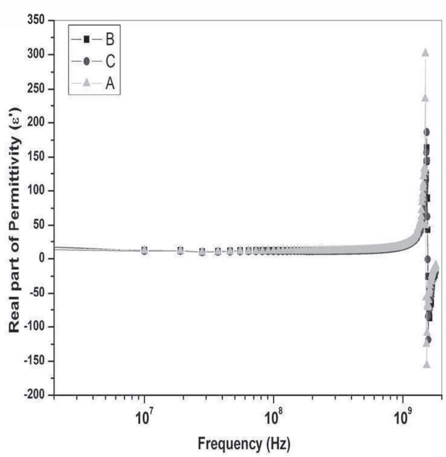


Figure 3. Variation of Real part of permittivity with frequency for composites

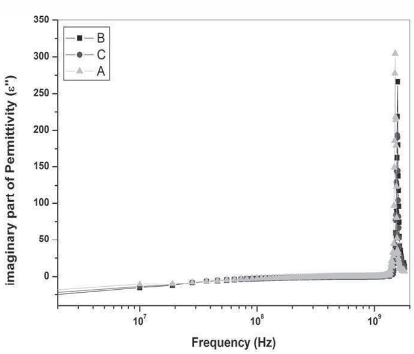


Figure 4. Variation of imaginary part of permittivity with frequency for composites

The frequency dependence of imaginary part of permittivity (ε″) has been measured on all the composites in the range of 1MHz to 1.8GHz and obtained results were plotted in Figure 4. A figure suggests that the value of ε″ is small and remains almost constant from 1 MHz to 600 MHz. The ε″ value increases with an increase of frequency from 600 MHz and finally shows a resonance peak above 1.4 GHz. The variation of ε″ with frequency may be explained similar to that of ε' variation with frequency. These relative permittivity and imaginary part of permittivity factors is very important for Multi-Layer Chip Inductor (MLCI) where ferrite thin films are separated by electrodes and the parasitic capacitance of this assembly is very important for multilayer devices. The capacitance and tan δ/ imaginary part of permittivity should be as low as possible for a good inductor material. All the composite ferrites shows very low capacitance and loss factor (tan δ) for wide frequency range.

CONCLUSION

High dense (90%), homogeneous and small grained (1-x) $Ni_{0.53}Cu_{0.12}Zn_{0.35}Fe_{1.88}O_4$ + (x) $Gd_{0.2}Ce_{0.8}O_3$(x=0.10,0.20,0.30) were prepared using conventional sintering method at 1000°C/1200min. The average grain sizes of all the composite lies between 90nm and 120nm. The present composites show low dielectric constant and low losses at 1MHz and resonant frequency of all the sintered samples found to be greater than 1.4GHz. As the ferrite content increases in the composite materials, the ferroelectric transition temperature is shifting towards the lower frequency side.

ACKNOWLEDGEMENTS

The authors are thankful to UGC-CSIR and DST New Delhi for providing financial assistance.

The Author is also thankful to CVR college of Engineering, Ibrahimpatan for carrying out research work in the college.

REFERENCE

[1] R. J. Charles, A. R. Achuta, U. S. Patent No. 4966625 (1990).
[2] T. Krishnaveni, B. R. Kanth, V. S. R. Raju, S. R. Murthy, J. All. Compd. 414 (1-2) (2006) 282
[3] S.R. Murthy, J. Mater. Sci. Lett. 21 (2002) 657.
[4] S.Bharadwaj, T .Ramesh, K.Sadhana and S.R.Murthy, Proceedings of the "International Conference on Nanoscience, Engineering and Technology" (ICONSET 2011 by IEEE (978-1-4673-0073-5/11) 206, 2011-2012.
[5] L. Zhao, Y. Cui, H. Yang, L. Yu, W. Jin, S. Feng, Mat. Lett. 60 (2006) 104.
[6] M. A. Ahmed, E. Ateia, F. M. Salem, Physica B (2008) (Accepted).
[7] G. L. Sun, J. B. Li, J. J. Sun, X. Z. Yang, J. Magn. Magn. Mater. 281 (2004) 173.
[8] J. Sun, J. Li, G. Sun, J. Magn. Magn. Mater. 250 (2002) 20.
[9] S. Solyman, Cera. Intern. 32 (2006) 755
[10] A. A. Sattar, K. M. El-Shokrofy, J. Phys. IV C1 (1997) 245
[11] M. A. Ahmed, N. Okasha, M. M. E. Sayed, Cera. Intern. 33 (1) (2007) 49.
[12] Z. Yue, Shaofeng Chen and Xiwei Qi, J.Alloys.Comp Vol 375 (2004) 243.

Minimization of Backbone Nodes in Wireless Mobile Backbone Networks using Approximation Algorithms

Dr. B. B. Jayasingh[1] and Sumitra Mallick[2]
[1]CVR College of Engineering, Department of IT, Ibrahimpatan, R.R.District, A.P., India
Email: bbjayasingh9@rediffmail.com
[2] Guru Nanak Institute of Technology, Department of IT, Ibrahimpatan, R.R.District, A.P., India
Email: sumitra_mallick2008@rediffmail.com

***Abstract*—We study a novel hierarchical wireless networking approach in which some of the nodes are more capable than others. In such networks, the more capable nodes can serve as Mobile Backbone Nodes and provide a backbone over which end-to-end communication can take place. Our approach consists of controlling the mobility of the Backbone Nodes in order to maintain connectivity. We formulate the problem of minimizing the number of backbone nodes and refer to it as the Connected Disk Cover (CDC) problem. We show that it can be decomposed into the Geometric Disk Cover (GDC) problem and the Steiner Tree Problem with Minimum Number of Steiner Points (STP-MSP). We prove that if these sub problems are solved separately by γ- and δ–approximation algorithms, the approximation ratio of the joint solution is $\gamma+\delta$. Then, we focus on the two sub problems and present a number of distributed approximation algorithms that maintain a solution to the GDC problem under mobility. We show that this approach can be extended in order to obtain a joint approximate solution to the CDC problem.**

***Index Terms*—Approximation algorithms, controlled mobility, distributed algorithms, disk cover, wireless networks.**

I. Introduction

WIRELESS Sensor Networks (WSNs) and Mobile Ad Hoc Networks (MANETs) can operate without any physical infrastructure (e.g., base stations). Yet, it has been shown that it is sometimes desirable to construct a *virtual backbone* on which most of the multi-hop traffic will be routed [6]. If all nodes have similar communication capabilities and similar limited energy resources, the virtual backbone may pose several challenges. For example, bottleneck formation along the backbone may affect the available bandwidth and the lifetime of the backbone nodes. In addition, the virtual backbone cannot deal with network partitions resulting from the spatial distribution and mobility of the nodes. Alternatively, if some of the nodes are more capable than others, these nodes can be dedicated to providing a backbone over which reliable end-to-end communication can take place.

A novel hierarchical approach for a *Mobile Backbone Network* operating in such a way was recently proposed and studied by Rubin *et al.* (see [7, 9] and references therein) and by Gerla *et al.* (e.g., [8]). We develop and analyze novel algorithms for the construction and maintenance (under node mobility) of a Mobile Backbone Network. We focus on *controlling the mobility* of the more capable nodes in order to maintain network connectivity and to provide a backbone for reliable communication.

A Mobile Backbone Network is composed of two types of nodes. The first type includes static or mobile nodes (e.g., sensors or MANET nodes) with limited capabilities. We refer to them as *Regular Nodes* (RNs). The second type includes mobile nodes with superior communication, mobility, and computation capabilities as well as greater energy resources (e.g., Unmanned-Aerial-Vehicles).We refer to them as *Mobile Backbone Nodes* (MBNs). The main purpose of the MBNs is to provide a mobile infrastructure facilitating network-wide communication. We specifically focus on minimizing the number of MBNs needed for connectivity. Yet, the construction of a Mobile Backbone Network can improve other aspects of the network performance, including node lifetime and Quality of Service as well as network reliability and survivability.

The set of MBNs has to be placed such that (i) every RN can directly communicate with at least one MBN, and (ii) the network formed by the MBNs is connected. We assume a *disk* connectivity model, whereby two nodes can communicate if and only if they are within a certain communication range. We also assume that the communication range of the MBNs is significantly larger than the communication range of the RNs.

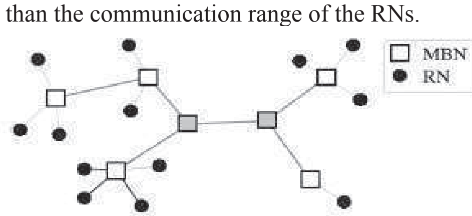


Figure 1. Illustrates an example of the architecture of a Mobile

We term the problem of placing the *minimum* number of MBNs such that both of the above conditions are satisfied as the *Connected Disk Cover (CDC)* problem. While related problems have been studied in the past [2], [6], [3] this papers are one of the first attempts to deal with the CDC problem. Our first approach is based on a framework that *decomposes* the CDC problem into two sub problems. We view the CDC problem as a two-tiered problem. In the first phase, the minimum number of MBNs such that all RNs are *covered* (i.e., all RNs can communicate with at least one MBN) is placed. We refer to these MBNs as *Cover MBNs* and denote them in Fig. 1 by white squares. In the second phase, the minimum number of MBNs such that the MBNs' network is connected is placed. We refer to them as *Relay MBNs* and denote them in Fig. 1 by gray squares. In the first phase, the Geometric Disk Cover (GDC) problem [1, 3] has to be solved, while in the second phase, a Steiner Tree Problem with Minimum Number of Steiner Points (STP-MSP) [4, 5, 10] has to be solved. We show that if these sub problems are solved separately by γ- and δ–approximation algorithms, the approximation ratio of the joint solution is γ+δ.

We develop a number of practically implementable distributed algorithms for covering mobile RNs by MBNs. We assume that all nodes can detect their position via GPS or a localization mechanism. This assumption allows us to take advantage of location information in designing distributed algorithms. We obtain the worst case approximation ratios of the developed algorithms and the average case approximation ratios for two of the algorithms. Finally, we evaluate the performance of the algorithms via simulation, and discuss the tradeoffs between the complexities and approximation ratios.

Our first main contribution is a decomposition result regarding the CDC problem. Other major contributions are the development and analysis of distributed approximation 1A δ approximation algorithm for a minimization problem always finds a solution with value at most δ times the value of the optimal solution. Algorithms for the GDC problem in a mobile environment, as well as the design of a novel Discretization Approach for the solution of the STP-MSP and the CDC problem.

II. PROBLEM FORMULATION

We consider a set of *Regular Nodes (RNs) distributed in the* plane and assume that a set of *Mobile Backbone Nodes (MBNs)* has to be deployed in the plane. We denote by N the collection of Regular Nodes {1,2,...n} , by $M=\{d_1, d_2, \ldots d_n\}$ collection of MBNs, and by d_{ij} the distance between nodes i and j. An RN_i can communicate bi-directionally with another node $_j$ (i.e., an MBN) if the distance between i and j, d_{ij} <=r, we denote by D=2r the diameter of the disk covered by an MBN communicating with RNs. Regarding the MBNs, we assume that MBN_i can communicate with MBN_j if d_{ij} <= R, where R>r.

We assume that the RNs and MBNs have both a communication channel (e.g., for data) and a low-rate control channel. For the communication channel, we assume the disk connectivity model. Namely, an RN can communicate bi-directionally with another node (i.e., an MBN) if the distance between i and j, d_{ij} <=r. We denote by D=2r the diameter of the disk covered by an MBN communicating with RNs. Regarding the MBNs, we assume that MBN i can communicate with MBN j if d_{ij} <=R, where R>r. For the control channel, we assume that both RNs and MBNs can communicate over a much longer range than their respective data channels. Since given a fixed transmission power, the communication range is inversely related to data rate, this is a valid assumption.

A. Connected Disk Cover (CDC) problem

Given a set of RNs (N) distributed in the plane, place the smallest set of MBNs such that:

1) For every RN_i ε N there exists at least one MBN $_j$ ε M such that d_{ij} <=r
2) The undirected graph G = (M,E) imposed on M (i.e., $\forall_{k,l}$ M, define an edge (k, l) ε E if d_{kl} <=R) is connected.

The RNs are mobile and some of the MBNs move around in order to maintain a solution the CDC problem. We will study both the case in which the nodes are static, and the case in which the RNs are mobile and some of the MBNs move around in order to maintain a solution the CDC problem. We assume that there exists some sort of MBN routing algorithm, which routes specific MBNs from their old locations to their new ones.

We propose to solve the CDC problem by decomposing it into two NP-Complete sub problems: the Geometric Disk Cover (GDC) problem and the Steiner Tree Problem with Minimum number of Steiner Points (STP-MSP).

B. Geometric Disk Cover (GDC)

Given a set N of RNs distributed in the plane, place the smallest set M of cover MBNs such that:

1) For every RN_i ε N there exists at least one MBN $_j$ ε M such that d_{ij} <=r
2) A set of cover MBNs is given and there is a need to place the minimum number of relay MBNs such that formed network is connected.

The second sub problem deals with a situation in which a set of *Cover MBNs* is given and there is a need to place the minimum number of *Relay MBNs* such that the formed network is connected (i.e., satisfies only property (2) in the CDC problem definition). This sub problem is equivalent to the Steiner Tree Problem with Minimum Number of Steiner Points (STP-MSP) [10].

C. Steiner Tree Problem (STP)

Given a set of cover MBNs (M_{cover}) distributed in the plane, place the smallest set of Relay MBNs (M_{relay}) such that the undirected graph G = (M,E) imposed on M = M_{cover} U M_{relay}(i.e., $\forall$ $_{k,l}$ M, define an edge (k, l) ε E if d_{kl} <=R) is connected.

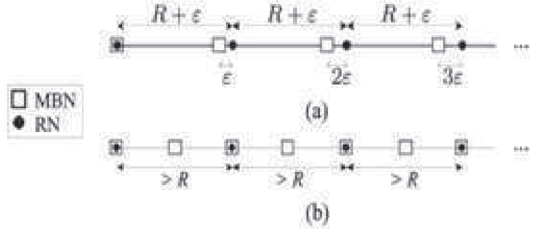


Figure 2. Tight example of the approximation ratio of the decomposition algorithm (a) optimal solution and (b) decomposition algorithm solution.

A tight example of this fact is illustrated in Fig. 2. Fig. 2(a) shows an node instance of the CDC problem, where ε<<r refers to a sufficiently small constant. Also shown is the optimal solution with cost n MBNs. Fig. 2(b) shows a potential solution obtained by using the decomposition framework (with γ=δ=1), composed of an optimal disk cover and an optimal STP-MSP solution. The cost is n+n-1=2n-1MBNs. This example highlights the facts that under the Decomposition Framework, the cover MBNs are placed without considering the related problem of placing the relay MBNs.

III. Placing The Cover MBNS—Strip Cover

Hochbaum and Maass[5, 11] introduced a method for approaching the GDC problem by (i) dividing the plane into equal width strips, (ii) solving the problem locally on the points within each strip, and (iii) taking the overall solution as the union of all local solutions. Below we present algorithms that are based on this method. These algorithms are actually two different versions of a single generic algorithm. The first version locally covers the strip with rectangles encapsulated in disks while the second version locally covers the strip directly with disks. We then generalize (to arbitrary strip widths) the effects of solving the problem locally in strips and use this extension to provide approximation guarantees. Finally, we discuss distributed implementations of these algorithms.

3.1 Centralized Algorithms

For simplicity of the presentation, we start by describing the centralized algorithms. The two versions of the Strip Cover algorithm (*Strip Cover with Rectangles*—SCR and *Strip Cover with Disks*—SCD) appear below. In line 6, the first version (SCR) calls the *Rectangles* procedure and the second one (SCD) calls the *Disks* procedure. *The input is a set of points (RNs) N={1, 2, ...n}* and their (x, y) coordinates, (i_x, i_y) $\forall$ i. The output includes a set of disks (MBNs) M={d_1, d_2, …d_n} and their locations such that all points are covered. The first step of the algorithm is to divide the plane into K strips of width αD (D=2r). We denote the strips by S_j and the set of MBNs in strip by Ms_j. Figure 3. shows an example of the SCR algorithm and in particular of step 9 in which disks are placed such that they compactly cover all points in the rectangular area with -coordinate range.

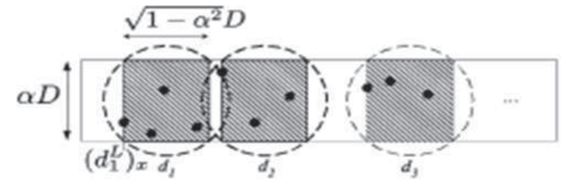


Figure 3. Illustrating step 9 of SCR algorithm

Algorithm 1: Strip Cover with Rectangles/Disks (SCR/SCD):

1. Divide the plane into K strips of width q_{sc} =αD
2. Ms_j ← ∅, $\forall_j$ = 1,…,K
3. For all strips S_j, j = 1,…,K do
4. While there exist uncoverd RNs in S_j
5. Let I be the leftmost uncovered RN in S_j
6. Call Rectangles(i) or call Disks(i)
7. Ms_j ← Ms_j ∪ d_k
8. Return $\cup_j$ Ms_j

Procedure Rectangles(i):

9. Place an MBN d_k such that it covers all RNs in the rectangular area with x-coordinates [i_x,i_x+ $\sqrt{1-\alpha^2}$D]
10. Return d_k

Procedure Disks(i):

11. Pd_k ← ∅ {set of RNs covered by the current MBN d_k}
12. While Pd_k ∪ i coverable by a single MBN (disk) do
13. Pd_k ← Pd_k ∪ i
14. If there are no more RNs in the strip then
15. Break
16. Let i be the nest leftmost uncovered RN in S_j not currently in Pd_k
17. Place MBN (disk) d_k, such that it covers the RNs Pd_k
18. Return d_k

3.3 Distributed Implementation

The SCR and SCD algorithms can be easily implemented in a distributed manner. The algorithms are executed at the RNs and operate within the strips. The SCR algorithm executed at an RN is described below. Recall that we denote the RNs within a strip according to their order from the left (i.e., i<j if i_x <= j_x).Ties are broken by node ID. Every RN that has no left neighbors within distance D initiates the disk placement procedure that propagates along the strip. The propagation stops once there is a gap between nodes of at least. If an RN arrives from a neighboring strip or leaves its MBN's coverage area, it initiates the disk placement procedure that may trigger an update of the MBN's locations within the strip. Notice that MBNs only move when a recalculation is required. Although the responsibility to place and move MBNs is with the RNs, simple enhancements would allow the MBNs to reposition themselves during the maintenance phase. If after a recalculation, an MBN is not repositioned, then it is not required and can be used elsewhere. The time complexity (i.e., number of rounds) is O(n). The computation complexity is Olog n). Control information has to be transmitted between RNs over a Distance D=2r.

Algorithm 2: Distributed SCR (at RN i)

Initialization:

1. Let G_i be the set of RNs j such that j<i and $i_x - j_x \le D$
2. If $G_i = \emptyset$ then
3. Call place MBN

Construction and Maintenance:

4. If MBN placed message received then
5. Call place MBN
6. If i is disconnected from its MBN or enters from a neighboring strip then
7. If there is at least one MBN within distance r then
8. Join one of these MBNs
9. Else call place MBN

Procedure place MBN

10. Let i^R be the rightmost RN s.t $(i^R)_x \le i_x + \sqrt{1-\alpha^2}D$
11. Place MBN d_k covering RNs j, where $j_x \in [i_x, (i^R)_x]$
12. If $(i^R+1)_x - (i^R)_x \le D$ then
13. Send an MBN placed message to i^R+1

Algorithm 3: Simple 1-D [13] with $\sqrt{1-\alpha^2}D = 2/3$

1. Initialize the cover greedily {using the SCR algorithm}
2. Maintain the leftmost RN and rightmost RN of each MBN rectangle
3. If two adjacent MBN rectangles come into contact then exchange their outermost RNs
4. If a set of RNs covered by an MBN becomes too long {the separation between its leftmost ad rightmost RNs become greater than 2/3} then
   Split off its rightmost RN into a singleton MBN
   Check whether rule 4 applies
5. If two adjacent MBN rectangles fit in a 2/3 rectangle then
   Merge the two MBNs

CONCLUSION

The architecture of a hierarchical Mobile Backbone Network has been presented only recently. Such a design can significantly improve the performance, lifetime, and reliability of MANETs and WSNs. We concentrate on placing and mobilizing backbone nodes, dedicated to maintaining connectivity of the regular nodes. We formulated the Backbone Node Placement problem as the Connected Disk Cover problem and showed it can be decomposed into two sub problems. A new approach for the solution of the second sub problem (STP-MSP) and of the joint problem (CDC) has also been discussed. We showed that when it is used to solve the CDC problem in a centralized manner, the number of the required MBNs is significantly reduced.

REFERENCES

[1] Anand Srinivas*, Gil Zussman, and Eytan Modiano,* Construction and Maintenance of Wireless Mobile Backbone Networks, IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 17, NO. 1, FEBRUARY 2009, pp. 239-252.

[2] D. Chen, D.-Z. Du, X.-D. Hu, G.-H. Lin, L. Wang, and G. Xue, "Approximations for Steiner trees with minimum number of Steiner points," *J. Global Optimization, vol. 18, no. 1, pp. 17–33, 2000.*

[3] M. Franceschetti, M. Cook, and J. Bruck, "A Geometric Theorem for Approximate Disk Covering Algorithms," Caltech, Technical Report ETR035, 2001.

[4] T. Gonzalez, "Covering a set of points in multidimensional space," *Inf. Proc. Lett., vol. 40, no. 4, pp. 181–188, 1991.*

[5] G. Lin and G. Xue, "Steiner tree problem with minimum number of Steiner points and bounded edge-length," *Inf. Proc. Lett., vol. 69, no.* 2, pp. 53–57, 1999.

[6] B. Das, R. Sivakumar, and V. Bharghavan, "Routing in ad-hoc networks using a virtual backbone," in *Proc. IEEE ICCCN'97*, Sep. 1997.

[7] I. Rubin, A. Behzad, R. Zhang, H. Luo, and E. Caballero, "TBONE: A mobile-backbone protocol for ad hoc wireless networks," in *Proc. IEEE Aerospace Conf.*, Mar. 2002.

[8] D. Gu, G. Pei, M. Gerla, and X. Hong, "Integrated hierarchical routing for heterogeneous multi-hop networks," in *Proc. IEEE MILCOM'00*, 2000.

[9] J. Hershberger, "Smooth kinetic maintenance of clusters," in *Proc. ACM SoCG'03*, Jun. 2003.

[10] J. S. Provan, "Convexity and the Steiner tree problem," *Networks*, vol. 18, no. 1, pp. 55–72, 1988.

[11] P. Klein and R. Ravi, "A nearly best-possible approximation algorithm for node-weighted Steiner trees," *J. Algorithms*, vol. 19, no. 1, pp. 104–115, 1995.

A Study on Autonomic Cloud Environment for Hosting Electronic Health Records

S. Jyothsna
CVR College of Engineering, Department of IT, Ibrahimpatan, R.R.District, A.P., India
Email: jyothsna.sundaragiri@gmail.com

Abstract—**Cloud computing promises a more cost effective enabling technology to outsource storage and computations. Healthcare systems that are truly open, scalable, heterogeneous and capable of supporting the personnel. The use of cloud computing in healthcare will span both clinical and non-clinical applications. Electronic Health Records (EHR) offer patients the opportunity to access their own medical records. Apart from EHR, cloud computing will also apply to other clinical uses, such as physician order entry and software imaging and pharmacy use. Non-clinical uses will likewise benefit, including management apps for patient billing, claims and revenue cycle management.**

***Index Terms*—Cloud Computing, HealthCare, EHR, EMR, HITECH, Google Health, Microsoft Health Vault, CloudSim.**

I. Introduction

1.1.1 Cloud Computing

The technical foundations of Cloud Computing include Service-Oriented Architecture (SOA) and Virtualizations of hardware and software. The goal of Cloud Computing is to share resources among the cloud service consumers, cloud partners, and cloud vendors in the cloud value chain. Cloud Computing is likely to benefit a number of sectors and health being one of them.

1.1.2 Health care in Cloud Computing

For all the innovations transforming the healthcare industry, one area where it remains almost universally behind the times is in the use of information technology. Even some of the best equipped hospitals still rely on telecommunication as the primary communications tool for doctors; coordinating care schedules and other administrative processes remain cumbersome. At the same time, consumers and patients have very little transparency on health plan costs and covered services, with health insurance websites providing very little clarity to alleviate the problem.

One reason for this slow adoption of installing advanced IT solutions is attributed to high equipment costs, which usually involves new servers, storage and applications. In addition, there are high maintenance costs associated with keeping these systems up-and-running with software patches and upgrades. As a result, hospital administrators and boards would rather invest their limited financial budgets on new medical equipment or hire additional doctors, specialists and nurses. For smaller private clinics and doctors, the cost of technology is simply too high to even consider the option.

There is a way to eliminate the high capital expenses while still acquiring the latest IT solutions to improve healthcare services using cloud computing.

Imagine reading your electronic health record on your smart phone, or better yet consulting a doctor's opinion live from your tablet! These are great possible services enabled by cloud computing applications, which will change the nature of the competition between healthcare companies. To this end, like in any other field, CEOs of the health industry understand the paramount importance of cloud in their business.

Healthcare organizations will continue to see IT budgets being squeezed. CIOs and IT managers need to find creative approaches to delivering higher levels of service to their users. Cloud computing offers a viable alternative for many organizations of all sizes and health professionals seeking to solve current problems while delivering the best possible healthcare service.

Cloud computing can help clinicians and hospitals to coordinate and exchange information more efficiently. The use of cloud computing architecture helps in eliminating the time and efforts, needed to roll a healthcare IT application in a hospital.

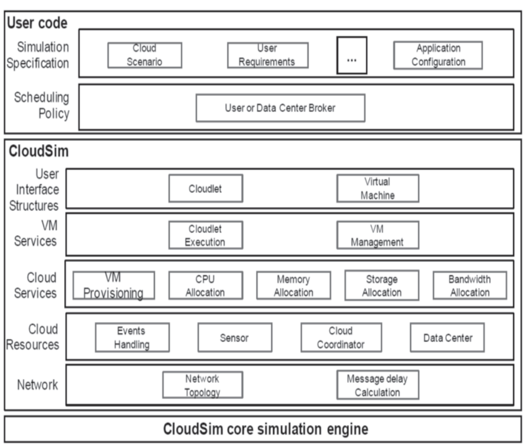
1.1.3 Electronic Health Records

An electronic health record (EHR) is an evolving concept defined as a systematic collection of electronic health information about individual patients or populations. It is a record in digital format that is theoretically capable of being shared across different health care settings. In some cases this sharing can occur by way of network-connected enterprise-wide information systems and other information networks or exchanges. EHRs may include a range of data, including demographics, medical history, medication and allergies, immunization status, laboratory test results, radiology images, vital signs, personal stats like age and weight, and billing information. This transition provides numerous benefits for patients, health insurance companies and medical staff. For example, patients gain detailed access to their records, which they can share with family members or the doctors of their choice, and insurance companies save money by avoiding repetition in patient care.

When lives hang in the balance and speedy treatment is required, cloud based EHR and cloud based EMR can provide fast and accurate access to patient medical information. Many times, doctors and ER physicians simply cannot wait for a manual search of patient information before beginning treatment. This waiting time can be dramatically reduced through the use of EMR systems and EHR systems. Cloud based systems offer the fastest, safest and most secure medical information available to any medical office, clinic or hospital.

1.1.4 CloudSim

CloudSim goal is to provide a generalized and extensible simulation framework that enables modeling, simulation, and experimentation of emerging Cloud computing infrastructures and application services, allowing its users to focus on specific system design issues that they want to investigate, without getting concerned about the low level details related to Cloud-based infrastructures and services support for modeling and simulation of large scale Cloud computing data centers.



1.1.4.1 CloudSim Features:

- support for modeling and simulation of virtualized server hosts, with customizable policies for provisioning host resources to virtual machines
- support for modeling and simulation of energy-aware computational resources
- support for modeling and simulation of data center network topologies and message-passing applications
- support for modeling and simulation of federated clouds
- support for dynamic insertion of simulation elements, stop and resume of simulation
- support for user-defined policies for allocation of hosts to virtual machines and policies for allocation of host resources to virtual machines

II. Related Work

Healthcare providers looking at automating processes at lower cost and higher gains. Cloud computing can act as an ideal platform in the healthcare IT space. A number of hospitals could share infrastructure with large number of systems linked together. By this pooling the hospitals automatically reduce the cost and increase utilization. The resources are delivered only when they are required. This also means real-time availability of patient information for doctors, nursing staff and other support services personnel from any internet enabled device.

2.1 2009 HITECH Act

The Health Information Technology for Economic and Clinical Health Act set meaningful use of interoperable EHR adoption in the health care system as a critical national goal and incentivized EHR adoption.The "goal is not adoption alone but 'meaningful use' of EHRs — that is, their use by providers to achieve significant improvements in care.

2.2 Google Health

Google Health is based on open standards (Continuity of Care Record for data exchange, SOAP for the web-services interoperability), and provides a development API, programming libraries and test infrastructure. Google guarantees it will protect the privacy of the information by giving the user complete control over it. To this end, Google Health features no advertising. Google Health is oriented towards the U.S market, as the third-party services it uses are exclusively American.

2.3 Health Vault

Health Vault consists of two distinct products – an electronic repository for health data and a specialized search engine for health information on the World Wide Web, both free to users. Health Vault is sometimes described as “PayPal for health information” for being able to store and share medical information at the discretion of its owner, as well as for utilizing similar security features. Health Vault stands out from other EHR providers because of its extensive partner network, particularly in the area of EHR.

III. System Design

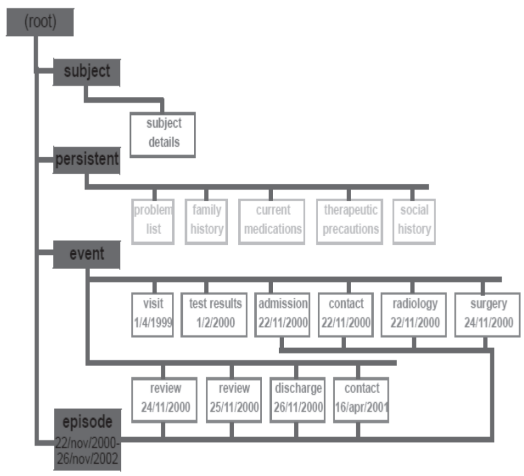
3.1 cloud based EHR system

A cloud based EHR system works in the following ways:

- ➢ Relies on the Internet for record storage and retrieval

- Electronic transmission of patient health records and medical information
- Medical information is accessed using standard Internet applications
- Access on any PC, laptop or smartphone

3.1.1 EHR Reference Model:

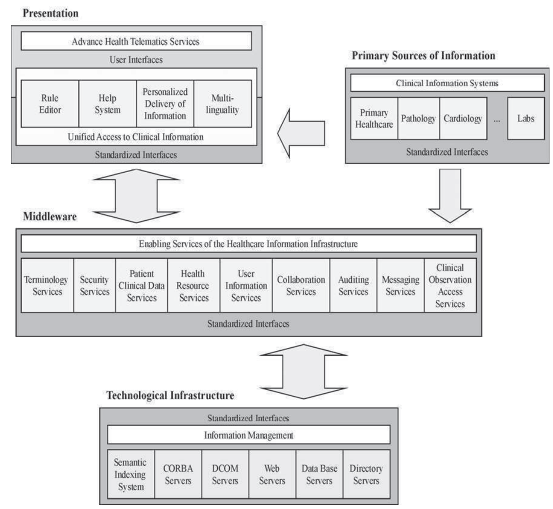


Subject: A composition containing clinically relevant demographic data of the patient.
Persistent: Compositions containing information which is valid in the long term.
Event: Compositions containing information whose currency is limited to the short term after the time of committal.
Episode_xxx: rather than using a single 'event' folder, it may be convenient to group event compositions into episodes (periods of treatment at a health care facility relating to one or more identified problems) and/or other categories such as on the basis of type of healthcare (orthodox, homeopathic, etc).

A justification for these particular categories is based on patterns of access. The persistent category consists of a dozen or so compositions described above, and which are continually required by querying (particularly lifestyle, current problems and medications). The event category consists of clinical data whose relevancy fades fairly quickly, including most measurements made on the patients or in pathology. Compositions in this category are thus potentially very numerous over the patient's lifetime, but of decreasing relevance to the clinical care of the patient in time; it therefore makes sense to separate them.

A patient's EHRs are typically dispersed over a wide range of distributed EMR systems in clouds. Different EMR systems have different data schemas to manage logical and semantic relationships between data elements drawn from various medical domains. Such medical domains include patient demographics, labs, medications, encounters, imaging and pathology reports,and a variety of other medical domains from primary, speciality and acute care settings. Each node in the hierarchical structure is labeled and the root of the hierarchical structure represents a particular EHR instance.

3.2 System Architecture



The above architecture is divided into

- Presentation-This Presents the Required information most effectively.
- Primary sources of information-where the key information about EHR is placed.
- Middleware-This is the layer which consists of various services like technology services, security services, collaboration services, patient clinical data services, user information services etc..
- Technological infrastructure-This consists of various Distributing Computing technologies like CORBA, DCOM etc.. for sharing and processing Electronic heath records.

3.3 EHR Services Offered

- Infrastructure Hosting
- Software as a Service
- Desktop virtualization
- Unlimited storage capacity
- Managed backup and server maintenance
- Top tier security & reliability
- IT Consulting
- Remote connectivity
- Automated calling programs
- Interactive voice response (IVR) technology
- VoIP

IV. Comparative Analysis

4.1 Cloud-Based EHR vs. Client-Server

EHR systems basically fall into two categories: cloud-based or client-server. In a cloud-based system, a practice's data is stored on external servers and can be accessed via the web, requiring only a computer with an Internet connection.

Client-server systems store data in house, requiring a server, hardware and software be installed in the physician's office. While in-house servers have traditionally been the norm, practices are increasingly switching to the cloud for a number of reasons.

4.1.1 Benefits of Cloud-Based EHR

1. Implementation is much simpler with cloud-based EHR systems: EHR software runs on the web instead of the computer, meaning no hardware or software installation. Practices can prevent interruption of cash flow and get a faster return on investment with an implementation process much quicker than traditional client-server systems.
2. Practices realize tremendous savings from cloud-based EHR systems: One of the largest hurdles for small medical practices is the initial cost of EHR installation. Client-server systems can cost $40,000 or more just to get set up, and then the licensing fees, maintenance costs, updates and patches cost more on top of that.

Since cloud-based EHR requires no hardware installation or software licenses, implementation is a fraction of the cost. Practices pay a monthly fee, like a utility bill, as part of an arrangement called software as a service (SaaS).

3. IT resource requirements are significantly reduced: when practices choose to move medical records to the cloud. Instead of requiring a team of IT experts to install, configure, test, run, secure and update hardware and software, all of that is done internally in the cloud by the SaaS provider. Updates are also done automatically in web-based systems, so practices are running on the most up-to-date version available.
4. Web-based software provides superior accessibility and collaboration: Over client-server systems because users are able to securely log in to the system from anywhere they have Internet connection. The ability to access the system outside of the office allows physicians, staff and patients to collaborate more effectively in a secure environment and provide better continuity of care.
5. Scalability is simplified with cloud-based systems: Small practices are able to expand without the standard IT growing pains. A web-based EHR system makes it easy to add new users, doctors or locations. The flexibility of web-based software allows small practices to think big and grow without breaking the bank.

V. Conclusion

5.1 *Summary*

Joining up diverse and sometimes discipline-specific and culturally specific kinds of clinical information to compose a whole-person EHR that can safely, legally and usably replace paper records is a complex challenge. A Cloud-based solution for healthcare requirements can bring quantifiable raw cost savings, as well as a number of strategic benefits designed for the betterment of the health care systems in a much broader sense. "These include operational cost, cost restructuring, competency alignment, risk management, rapid scalability and deployment benefits."

In practice, Today, 41.8 per cent of a healthcare organization's IT budget is allocated to a traditional IT deployment, whereas in two years' time this will decrease to 35.4 per cent, a decline of six percentage points. With this change comes an increase in the percentage of the overall IT budget allocated to private and public Cloud services. For infrastructure suppliers, this signals a very real change in how firms will procure services and solutions in future.

References

[1] Shinji kikuchi,shelly suchdeva,subhash balla,"cloud computing based PHR architecture using multilayers model",November,2012.

[2] Ahmed.S,Abdullah.A,"E-healthcare and data management services in a cloud", Dec. 2011.

[3] Rolim,C.O, "A Cloud Computing Solution for Patient's Data Collection in Health Care Institutions", Feb,2010 .

[4] "Consumers Compare Online Personal Health Record (PHR) Applications", A White Paper by User Centric, Inc.February 2, 2009.

[5] "The Cloud Simulator-CloudSim.",from www.cloudbus.org/cloudsim/.

[6] Koala: "A Discrete-Event Simulation Model of Infrastructure Clouds", web article available at: http://www.nist.gov/itl/antd/upload/Koala.pdf.

[7] T Beale, S Heard,D Kalra, D Lloyd"The *open*EHR Reference Model-EHR Information Model**"** © 2003-2007 The *open*EHR Foundation.

[8] *Dipak Kalra1, David Ingram1,"*Electronic Health Records" Centre for Health Informatics and Multiprofessional Education, UCL, London, UK.

Prevailing Practice and Perception on Managerial Aspects of Community Based Rehabilitation Projects in India

Narendra Kumar[1] and Dr. M.S.Bhat[2]
[1]JNTU, Hyderabad, A.P., India
Email: narendra_rehabilitation@yahoo.co.in
[2]CVR College of Engineering, Department of Management, Ibrahimpatan, R.R.District, A.P., India
Email: bhattms@rediffmail.com

***Abstract* - The major change in strategy in rehabilitation for people with disabilities over the past few years has been the expansion of services in the community. This has slowly gathered the momentum and has developed into a differentiated programme called 'Community Based Rehabilitation' (CBR). Ideally, a community-based rehabilitation programme is built on an integrated and decentralized managerial approach, in which both the service providing agencies and the community have a role to play.**

A few aspects of management of community based rehabilitation were discussed in the research work of Shamrock, 2009; Evans et al., 2001; Andrew & Dominic 2004; Cheausuwantavee, 2005; Kuipers & Harknett, 2008; Andrew et al., 2009; Cornielje et al., etc.. Not much research work has been found addressing the issues and aspects related to the management of CBR directly and extensively. Hence present research has been taken up to study the community based rehabilitation implemented by various disability rehabilitation organizations, from managerial perspectives.

In the study, the efforts have been made to understand the CBR projects in the Indian context by studying the prevailing practices of Managerial Aspects of CBR and the perception of CBR managers on the effectiveness of these Managerial Aspects of CBR. The correlation between practices and perception has been studied by making use of regression analysis.

The important results of the study include the construction of an assessment tool to study the prevailing practice of Managerial Aspects of CBR and the perception of CBR managers on the effectiveness of Managerial Aspects of CBR and the regression equation to predict the prevailing practice of Managerial Aspects of CBR based on the perception score of CBR managers. The study also highlights the relevance of managerial aspects of CBR in Indian context.

***Index Terms* – Disability, Rehabilitation, Community Based Rehabilitation, Managerial Aspects, Non Government Organizations, India**

---

[1] Research Scholar, JNTU, Hyderabad
[2] Professor of Management Science (Retd.), School of Management Studies, JNTU, Hyderabad and Prof &Dean (Evaluation) CVR College of Engineering

I. INTRODUCTION

The major change in strategy of rehabilitation for persons with disabilities over the past 25 years has been the expansion of services in the community. This has slowly gathered the momentum and has developed into a differentiated programme called 'Community Based Rehabilitation' (CBR). Among the recent strategies adopted for the rehabilitation of all persons with disabilities, 'community based rehabilitation' is of a special interest. This strategy has been found to be a viable alternative for the large populations of people with disabilities belonging to rural areas, especially those who are unable to access ongoing conventional rehabilitation services (Thomas & Thomas, 2008)[1].

Community-based rehabilitation (CBR) emerged in the 1970s with the intention to deliver "low-tech rehabilitation services" particularly to the large number of disabled people living in developing countries (Boyce & Lysak, 2000)[2]. Helander E (2007)[3] explained that in the late 1980s with the emergence of human rights for people with disabilities, community based rehabilitation shifted towards a greater focus on people and community development. Recently, community based rehabilitation is defined as "A strategy within community development for rehabilitation, equalization of opportunities, and social inclusion for all children and adults with disabilities" (ILO et al., 2002)[4].

Helander E (2002)[5] explained that the management in the context of CBR consists of all the efforts to ensure smooth functioning of the programme. This may include planning, policy-making, training of personnel, provision of resources, programme implementation at all levels, monitoring and evaluation. In the conventional system, the management practices existed in CBR were centralized and often practiced from the top to downward. If government is involved, it might make plans, get involved in the training of personnel, and introduce these projects at the district and the community level. The governments usually see communities as passive recipients of services and in their benevolence they have organized the services. As a result, involvement of local people may not be as much as it should be. If Non-Government Organizations (NGOs) are the service providers, they normally limit themselves to managing their own center. As long as the rehabilitation services involve coordination between a few institutions, the system can easily be managed centrally. But, when communities are involved in service delivery, large resources are needed, both centrally and in the periphery. Central direction induces a certain degree of passivity in the population. People often view that they have to wait for the government to come and offer services for them. This becomes frustrating for the people. Even governments in most developing countries find it difficult to manage the implementation of most essential services.

II. Review Of Literature

The review of literature shows that the articles on community based rehabilitation covers wide range of aspects related to CBR in developing countries. The key aspects of CBR throw light on the issues like differences in rehabilitation of children, adults, and the old age, different needs and services for urban and rural populations, gender inequalities in rehabilitation, specific approaches to types of disabilities, types of rehabilitation, issues involving the setting up the services, locus of control etc. From the managerial perspective the research work on community based rehabilitation seems to be limited. Some aspects of management of community based rehabilitation were discussed in the research work of Shamrock, 2009[6]; Evans et al., 2001[7]; Andrew & Dominic 2004[8]; Cheausuwantavee, 2005[9]; Kuipers & Harknett, 2008[10]; Andrew et al., 2009[11]; Cornielje et al., 2008[12]; who talked about the project evaluation part of community based rehabilitation in various ways. The studies of Powell et al., 2002[13]; Vijayakumar et al., 2003[14]; Pupulin & Aldén, 2002[15] were related to the quality of life of people with disabilities and stakeholders participating in CBR projects directly or indirectly. Impact analysis of CBR was carried out in detail by Chappell & Johannsmeier, 2009[16]; Arne, 2006[17]; Boyce & Cote, 2009[18]; Stilwell et al., 1998[19]. Though the research work quoted above talks about management aspects of CBR in bits and pieces, hardly any study was found addressing managerial aspects of CBR directly and extensively. Hence, the present research was taken up to study the community based rehabilitation from managerial perspectives, in the Indian context.

III. Research Mehodology

The scope of this study is wide from a concept point of view, because it covers broader aspects of community based rehabilitation management. In the study, efforts were made to understand the CBR projects in Indian context by studying the prevailing practices of managerial aspects of CBR and the perception of CBR managers on the effectiveness of these managerial aspects of CBR.

The correlation between practices and perception has also been studied by making use of regression analysis to test the null hypothesis "prevailing practices of managerial aspects of CBR does not depends on the perception of CBR managers regarding effectiveness of these managerial aspects of CBR" while analyzing the results.

Sample unit for the study was the organizations implementing community based rehabilitation in India. By using stratified random sampling method of probability sample design, one hundred and twenty one organizations were chosen for the study. These one hundred and twenty one organizations has representation from each zone (strata) named as; North zone, South zone, East zone, West zone and Central zone as presented in Table I.

TABLE I.
ZONAL DISTRIBUTION OF SAMPLE

Zones	Frequency	Percent
East Zone	13	10.7
North Zone	24	19.8
South Zone	53	43.8
West Zone	25	20.7
Central Zone	6	5.0
Total	121	100

East zone includes Arunachal Pradesh, Assam, Bihar, Manipur, Meghalaya, Mizoram, Nagaland, Orissa, Sikkim, Tripura and West Bengal while the North zone includes Haryana, Himachal Pradesh, Jammu and Kashmir, Punjab, Rajasthan, Uttarakhand, Uttar Pradesh, Chandigarh and Delhi. The South zone represents Andhra Pradesh, Karnataka, Kerala, Tamil Nadu, Andaman and Nicobar Islands, Lakshadweep and Puducherry while the West zone represents Goa, Gujarat, Maharashtra, Dadra and Nagar Haveli and Daman and Diu. The Central zone includes the Chhattisgarh, Jharkhand and Madhya Pradesh.

To evaluate the prevailing practices of managerial aspects of CBR in India and the perception of CBR managers on the effectiveness of managerial aspects of CBR, an assessment tool consisting nineteen domains was developed. The 114 test items under nineteen domains of managerial aspects of CBR were derived from the research work of various authors viz. Rao, L.G, 2002[20], Kumar & Bhat, 2009[21], Silvia T, 2001[22], Kundu et al., 2007[23], Das K et al., 2009[24], Hjøllund L, 1983[25], Bontis N, 1989[26] etc.

In order to establish the validity of assessment tool, it was given to fifteen professionals working in the field to obtain their agreement on each item. All the 114 items of the assessment tool were fully agreed upon by all the professionals. Further to this, a pilot study was also conducted with twenty subjects to find out the contribution of each item towards the grand score. Pearson's coefficient of correlation 'r' of each item with total score was highly significant ($p<0.01$)

and hence, all the 114 items planned were retained for the study.

To collect the primary data through survey, constructed assessment tool was circulated to all 121 organizations along with a covering letter by mail, e-mail and also through personal visits. The 86 responses received (duly filled assessment tool in all aspects), were taken up for the analysis by using SPSS software package 10.1 version.

IV. Result & Discussion

The analysis of data collected form 86 sample organizations indicates that among the selected sample, majority of the organizations were from South zone and were established during 1981 – 2000. Majority of the organizations were implementing their CBR projects in sub urban areas and focusing on either all types of disabilities or on mental retardation. Majority of these organizations were recognized as either special school or HRD training centers. In most of the cases, their CBR projects were funded by government (either state or central) and the annual budget of their CBR projects falls between Rupees 5 Lakhs and 10 Lakhs. In most of these CBR projects, the staff strength was between 10 and 20 with the staff having qualification of either diploma or graduation. Of the CBR managers who responded to the assessment tool, majority of were males between 20 years and 40 years of age. These respondents were either diploma holders or postgraduates with work experience of 3 to 15 years.

A. Analysis of mean score

Average score for all the 86 participants was calculated for part B and part C of assessment tool. The distribution of mean score for all the nineteen domains for the sample is presented in table: 02. The mean score exhibited in table: 02 indicates that for part B of the assessment tool, the mean score is the lowest in domain “Strategic management” (1.33) and the highest in “Individual attention to persons with disabilities” (4.50). For part C of the assessment tool, the lowest mean (1.47) is for the domain “Collaborative working” and the highest (4.52) is for “Individual attention to persons with disabilities”.

In part B, out of nineteen domains, twelve domains viz. Strategic management, Collaborative working, Strategy and leadership, Building social and intellectual capital, Knowledge management, Total quality rehabilitation management, Project management, Operational effectiveness, Resource management, Client relationship management, Service values and meaning and Organizational culture and management have their mean score lower than the total scale mean score (2.54). Whereas seven domains viz. Finance management, Human resource management, Human resource development, Capabilities & ownership, Trust & motivation, The rights and needs of persons with disabilities and Individual attention to persons with disabilities have their mean score greater than the total scale mean score (2.54).

In part C, thirteen domains viz. Strategic management, Collaborative working, Strategy and leadership, Building social and intellectual capital, Knowledge management, Total quality rehabilitation management, Project management, Operational effectiveness, resource management, client relationship management, Finance management, Service values and meaning and Organizational culture and management have their mean score lower than the total scale mean score (2.56). Whereas six domains viz. Human resource management, Human resource development, Capabilities and ownership, Trust and motivation, The rights and needs of persons with disabilities and Individual attention to persons with disabilities have their mean score greater than the total scale mean score (2.56).

Table II.
DISTRIBUTION OF MEAN SCORE

Domains	Max. Score	Mean Score	
		Part B	Part C
Service values and meaning	05	2.45	2.50
The rights and needs of PWD	05	4.48	4.50
Individual attention to PWD	05	4.50	4.52
Human resource development	05	3.42	3.53
Resource management	05	2.44	2.51
Collaborative working	05	1.48	1.47
Strategy and leadership	05	1.48	1.51
Trust and motivation	05	3.63	3.50
Capabilities and ownership	05	3.48	3.50
Operational effectiveness	05	2.41	2.49
Human resource management	05	3.41	3.51
Project management	05	1.59	1.52
Total quality rehabilitation management	05	1.53	1.53
Client relationship management	05	2.45	2.49
Organizational culture and management	05	2.49	2.47
Finance management	05	2.57	2.49
Knowledge management	05	1.53	1.52
Strategic management	05	1.33	1.52
Building social & intellectual capital	05	1.52	1.53
Mean	05	2.54	2.56
Total Score	95	48.19	48.63

B. Relationship between mean score and organizational characteristics

The analysis of variance though one way ANOVA for prevailing practice of managerial aspects of CBR as dependent variables and organizational characteristic of CBR implementing agencies as independent variables indicate that the mean score obtained for prevailing practice of managerial aspects of CBR is not related to the organization’s zonal representation and period of establishment.

The same test statistics indicates that there is a significant relationship between mean score of prevailing practice of managerial aspects of CBR and the CBR project location, disability focused in CBR projects, organizational recognition, funding source, annual budget, staff strength, and higher qualification of CBR staff.

C. Relationship between mean score and individual characteristics

Considering the mean score obtained for manager's perception as dependent variable and the manager's characteristics as independent variables, the analysis of variance though one way ANOVA reflects that the perception of manager's on prevailing practice of managerial aspects of CBR is not influenced by the gender of CBR managers.

The perception of CBR manager of the effectiveness of prevailing practice of managerial aspects of CBR depends on the age of the CBR manager, qualification of CBR manager and the work experience of CBR manager.

D. Test of null hypothesis

"Prevailing practices of managerial aspects of CBR does not depends on the perception of CBR managers on the effectiveness of these managerial aspects of CBR"

The hypothesis here tries to find out the relationship between practices of managerial aspects of CBR and the perception of CBR managers on the effectiveness of these managerial aspects of CBR. The mean score for perception of effectiveness was treated as independent variable and the mean score for the prevailing practices was treated as dependent variable. In order to find out that the practices on managerial aspects of CBR depend on the perception of CBR managers regarding effectiveness of these managerial aspects of CBR, the following tests were used:

Pearson product-moment correlation coefficient: Pearson correlation test statistics in table: 03 indicate a strong and positive correlation (r=.754) between practices of managerial aspects and the perception regarding effectiveness of these aspects.

Table III.
CORRELATIONS BETWEEN PRACTICE & PERCEPTION

		Practice
Perception	Pearson Correlation	.745**
	Sig. (2-tailed)	.000
	N	86

** Correlation is significant at the 0.01 level (2-tailed).

This correlation indicates that if the score on perception regarding effectiveness of managerial aspects of CBR increases, it is more likely to get higher score on prevailing practice of managerial aspects of CBR. To establish this relationship further, linear regression analysis was performed with the collected data by 'forced entry' method.

Linear regression analysis: In the linear regression analysis, R square was a measure of the proportion of variation in the scores that is explained by the variables in the model. The closer to 1 the more strongly the variables explain the response. The closer to 0 the less strongly the variables explain the response. The Table IV. indicates positive and fairly strong relationship between predictor (perception) and dependent variable (practice) with R value 0.755 closure to 1 and the R Square 0.570 also closure to 1.

Table IV.
MODEL SUMMARY[b] FOR PREVAILING PRACTICE

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
Perception of CBR Managers	.755a	.570	.565	6.567

a. Predictors: (Constant), Perception
b. Dependent Variable: Practice

The value of R Square for the prevailing practice of managerial aspects of CBR is 0.570 meaning that the model composed by the test variable (manager's perception on effectiveness) account for 57%. This regression value is fairly high and indicates that the perception of CBR managers on the effectiveness of prevailing practice of managerial aspects of CBR has significant influence on the practices of managerial aspects of CBR. The model itself is highly significant (ANOVA p = 0.000) and therefore this model is a better explanation than using just mean values (Table V).

Table V.
ANOVA [b] ON PREVAILING PRACTICE

Model		Sum of Squares	df	Mean Square	F	Sig.
Perception of CBR Managers	Regression	4794.418	1	4794.418	111.180	.000a
	Residual	3622.338	84	43.123		
	Total	8416.756	85			

a. Predictors: (Constant), Perception
b. Dependent Variable: Practices

The model presented in Table VI., composed of perception and a constant, where the constant represents the percentage of practice if there was no information on perception. The beta value (column B) for the constant is -33.569. The beta value of perception is 31.049 with significant level 0.000.

Table VI.
COEFFICIENTS [a] OF PREVAILING PRACTICE

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	-33.569	10.170		-3.301	.001	-53.793	-13.344
Perception of CBR Managers	31.049	2.945	.755	10.544	.000	25.194	36.905

a. Dependent Variable: practice

This indicates the slope of the regression line by offering information about two aspects: one, there is a positive or negative regression; second, the change in the predicted value for each unit change in the parameter. Regression equation derived from above statistics is:

Practice = -33.569 + 31.049 (Perception)

The positive slope in Figure 1. represents that the higher the values of these perception higher the chances of practice. The residual plot (Figure 2.) shows a random scatter of the points (independence) with a constant spread (constant variance).

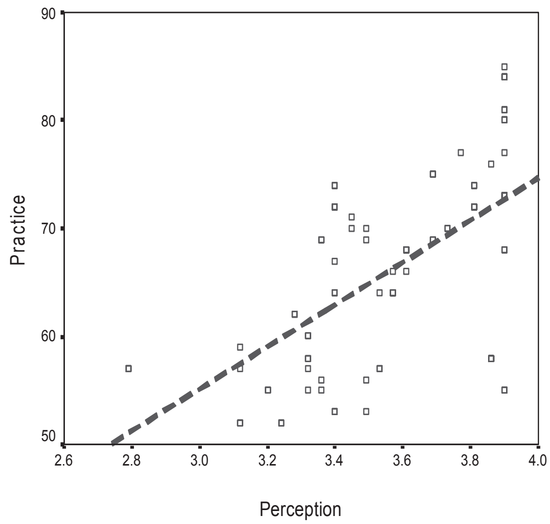


Figure 1. Slope of regression line (perception vs. Practice)

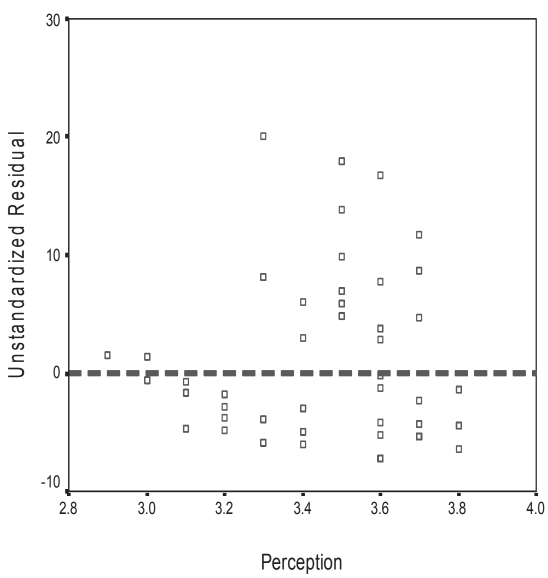


Figure 2. Residual plot

The standardized residual plot (Figure 3.) shows a random scatter of the points (independence) with a constant spread (constant variance) with some outliers (values beyond the ±2 standard deviation reference lines). These outliers are not treated or analyzed further in this study. The normal probability plot of the residuals (Figure 4.) shows the points close to a diagonal line; therefore, the residuals appear to be approximately normally distributed. Thus, the assumptions for regression analysis appear to be met.

Summary of regression analysis: At the 10% significance level, the data provide sufficient evidence to conclude that the slope of the population regression line is not 0 and, hence, perception of CBR managers on effectiveness of prevailing practice of managerial aspects of CBR is a useful predictor for prevailing practice of managerial aspects of CBR. Thus the null hypothesis “prevailing practices of managerial aspects of CBR does not depends on the perception of CBR managers on the effectiveness of these managerial aspects of CBR” is rejected.

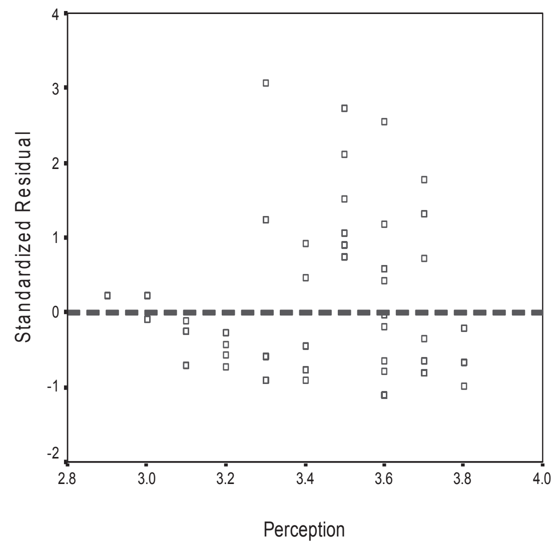


Figure 3. Standardized residual plot

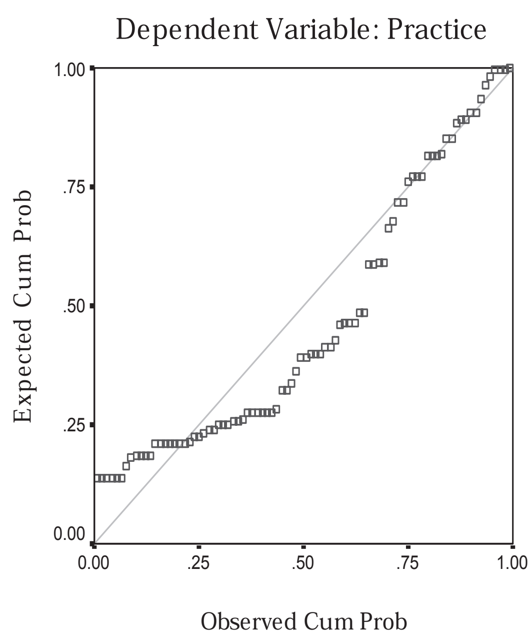


Figure 4. Normal P - P plot of regression standardized residual

Conclusions And Recommendations

The results of present study on prevailing practice and perception on managerial aspects of community based rehabilitation projects in India indicate that these organizations are rated as 'average' in terms of addressing the managerial aspects of CBR. The perception on effectiveness of managerial aspects of CBR is also rated as 'average' among CBR managers in India. The study points out that the prevailing practices of managerial aspects of CBR depends on the CBR project location, disability focused in CBR projects, organizational recognition, funding source, annual budget, staff strength, and higher qualification of CBR staff. The perception of CBR managers on effectiveness of managerial aspects of CBR depends on the age of the CBR manager, educational qualification of CBR manager and the work experience of CBR manager. Linear regression analysis indicates that there is a strong and positive correlation between practices of managerial aspects and the perception regarding effectiveness of these managerial aspects. Hence change in perception may lead to the change in practices of managerial aspects. In order to improve the management practices of CBR projects the perception of CBR managers need to be improved through a proper training and development in the area of managerial aspects of CBR.

Hence, the study suggests that managerial aspects of CBR cannot be ignored if a disability rehabilitation organization wants to implement CBR projects successfully. Top management need to put conscious efforts towards inclusion of managerial aspects of CBR in day to day affairs of CBR project management. Existing human resource involved with the CBR projects may also think of upgrading their skills to practice managerial aspects of CBR in day to day operations. Rehabilitation Council of India is the regulatory authority in India to develop and design the curriculum for developing human resource in the field of disability rehabilitation. Existing curriculum on CBR training does not have the adequate coverage on managerial aspects of CBR prescribed in this study. Inclusion of these aspects will not only improve the quality of existing long term training courses but also open up the scope of designing tailor made short term training programme in the field of Community Based Rehabilitation.

References

[1] Thomas M and Thomas M.J: Global trends in disability rehabilitation and their implications for leprosy programmes, Leprosy Review, 2008, Vol. 79, 10–16

[2] Boyce W and Lysak C: Community participation: Uncovering its meaning in CBR in Thomas M and Thomas MJ, editor's selected readings in community based rehabilitation: CBR in transition, Occasional publication of the Asia Pacific Disability Rehab Journal, Vol. 21 No. 1, 2000.

[3] Helander E: The origins of community based rehabilitation, Asia Pacific Disability Rehabilitation Journal, Vol. 18, No.2, 2007, 3-32.

[4] ILO, UNESCO, UNICEF and WHO: Community based rehabilitation (CBR) with and for people with disabilities, Joint Position Paper, unpublished paper, 2002.

[5] Helander E: Guide on the National Planning of Rehabilitation, Publications on Community-based Rehabilitation Management series - 1, United Nations Development Programme, Geneva Switzerland, 2002.

[6] Shamrock J: Evaluation of a pilot community based rehabilitation training programme in East Timor, Asia Pacific Disability Rehabilitation Journal, 2009, Vol.20 No.2, 67-78.

[7] Evans P.J, Zinkin P, Harpham T and Chaudury, G: Evaluation of medical rehabilitation in community based rehabilitation, Social science and medicine, 2001, Vol. 53, No. 3, 333-348.

[8] Andrew M. H. Siu and Dominic Y. Y. Chui: Evaluation of a community rehabilitation service for people with rheumatoid arthritis, Patient Education and Counseling, Vol. 55, No. 1, 2004, 128-140.

[9] Cheausuwantavee T: Community based rehabilitation in Thailand: current situation and development. Asia Pacific Disability Rehabilitation Journal. Vol. 16 No. 1, 2005.

[10] Kuipers P and Harknett S: Considerations in the quest for evidence in community based rehabilitation, Asia Pacific Disability Rehabilitation Journal, Vol.19, No.2, 2008.

[11] Andrew M.H. Siu, Daniel T.L. Shek and Peter K.K. Poon: Evidence-Based Research in Community Rehabilitation: Design Issues and Strategies, Hong Kong Journal of Occupational Therapy, Vol. 19, No. 1, 2009, 202-214.

[12] Cornielje H, Johan P Velema and Harry F: Community based rehabilitation programmes: monitoring and evaluation in order to measure results, Leprosy Review. Vol. 79, No. 1, 2008.

[13] Powell B.A, Mercer S.W, and Harte C: Measuring the impact of rehabilitation services on the quality of life of disabled people in Cambodja, Disasters, 2002, Vol. 26, No.2, 175-191.

[14] Vijayakumar V, Datta D, Karthika A, Thulasiraj RD, Nirmalan PK: Utilization of community-based rehabilitation services for incurably blind persons in a rural population of southern India, Indian Journal of Ophthalmology, Vol. 51, No.3, 2003.

[15] Pupulin E and Aldén M: Community-Based Rehabilitation as we have experienced it – voices of persons with disabilities, World Health Organization and Swedish Organizations of Disabled Persons International Aid Association, 2002

[16] Chappell P and Johannsmeier C: The impact of community based rehabilitation as implemented by community rehabilitation facilitators on people with disabilities, their families and communities within South Africa, Disability and Rehabilitation, Vol. 31, No.1, 2009.

[17] Arne H. Eide: Impact of Community-Based Rehabilitation Programmes: The Case of Palestine, Scandinavian Journal of Disability Research, Volume 8, No. 4, 2006, 84-92.

[18] Boyce W and Cote L: The influence of HIV/AIDS on community based rehabilitation in Dar Es Salaam,

Tanzania. Asia Pacific Disability Rehabilitation Journal, Vol. 20, No. 1, 2009, 52 – 72
[19] Stilwell P, Stilwell J, Hawley C and Davies C: Measuring outcome in community-based rehabilitation services for people who have suffered traumatic brain injury: the Community Outcome Scale. Clinical Rehabilitation, 1998, Vol. 12, No. 6, 54-62.
[20] Rao L.G: Human resource development in rehabilitation organizations – A study of select non-governmental organizations dealing with disabled persons in Andhra Pradesh, 2002, Ph.D. thesis submitted to Jawaharlal Nehru Technological University.
[21] Kumar and Bhat: Study on Organizational effectiveness of Disability Rehabilitation Organizations: Perception of Professional and Managers, Indian Journal of Training and Development, Vol. 39, No.4, 2009, 23-34, ISTD, New Delhi.
[22] Silvia T: Project Management Audit: Project Management Processes and Quality Checklist of International ISO 10006:1997(E), PRINCE2 and PMBOK Methodologies, 2001.
[23] Kundu C. Subhash, Malhan D, Kumar P: Human resource management practices in shipping companies, Delhi Business Review, Vol. 8, No. 1, 2007.
[24] Das K, Parmar J, Sadanand V Kumar: Customer Relationship Management (CRM) Best Practices and Customer Loyalty - A Study of Indian Retail Banking Sector, European Journal of Social Sciences, Volume 11, No.1, 2009, 28-35.
[25] Hjøllund L: Social Capital: A Standard Method of Measurement, Denmark, 1983.
[26] Bontis N: Intellectual Capital Questionnaire, Institute for Intellectual Capital Research, Canada, 1989.

ABOUT THE COLLEGE

CVR College of Engineering was established in 2001, and its eighth batch of students graduated from the College. CVR is ranked among the top few of more than 600 colleges in AP that started in the last decade and was also rated as the #1 co-educational college in Pass percentage among nearly 300 colleges under JNTU, Hyderabad for examinations held in May 2009, November 2009 and May 2010. It is the expectation of the academic community that CVR is on the successful path to be in the TOP-5 amongst all colleges in AP in the next few years.

The College was the first college in Osmania University area that was promoted by NRI technology professionals resident in the US. The NRI promoters are associated with cutting-edge technologies of the computer and electronics industry. They also have strong associations with other leading NRI professionals working for world-renowned companies like IBM, Intel, Cisco, Motorola, AT&T, Lucent and Nortel who have agreed to associate with the College with a vision and passion to make the College a state-of-the-art engineering institution.

The College has been given Permanent Affiliation and Autonomous Status by JNTUH.

CALL FOR PAPERS:

Papers in Engineering, Science and Management disciplines are invited for Publication in our Journal. Authors are requested to mail their contributions to Editor, CVR Journal of Science and Technology (Email Id:chalapatiraokv@gmail.com).

Papers are to be written using a Standard Template, which may be obtained on request from the Editor.

CVR JOURNAL OF SCIENCE & TECHNOLOGY

CVR COLLEGE OF ENGINEERING
(An Autonomous College affiliated to JNTU Hyderabad)
Mangalpalli (V), Ibrahimpatan (M)
RR District, AP-501510
http://cvr.ac.in