# PREVENTION OF BLACK HOLE ATTACK IN MANET USING GENETIC ALGORITHM

**Arshdeep Kaur[1], Mandeep Kaur[2]**

*[1]Student, [2]Asst Prof, Guru Kashi University Talwandi Sabo, Bathinda (PB)*

## ABSTRACT

*Recently, Wireless ad hoc networks became a hot research topic among researchers due to their flexibility and independence of network infrastructures, such as base stations. Network Early work in MANET research has mainly focused on developing an efficient routing mechanism in such a highly dynamic and resource-constrained network. At present, several efficient routing protocols have been proposed for MANET. Most of these protocols assume a trusted and cooperative environment. However, in the presence of malicious nodes, the networks are vulnerable to various kinds of attacks. In MANET, routing attacks are particularly serious. So, this proposed work tries to design and implements GA algorithm with Black hole attack and prevent the system for threat using this DSR protocol.*

***Keywords: Mobile Ad Hoc Network, Genetic Algorithm, Dynamic Source Routing Protocol***

## I. INTRODUCTION

The military tactical and other security-sensitive operations are still the main applications of ad hoc networks, although there is a trend to adopt ad hoc networks for commercial uses due to their unique properties [1]. However, similar to other networks, MANET also vulnerable to many security attacks. MANET not only inherits all the security threats faced in both wired and wireless networks, but it also introduces security attacks unique to itself. In MANET, security is a challenging issue due to the vulnerabilities that are associated with it [2].

Intrusion detection is therefore incorporated as a second line of defense in addition to key based authentication schemes. The ranges of attacks that can be mounted on MANETs are also wider than in case of conventional static networks. In mobile wireless networks there is no infrastructure as such and so it becomes even more difficult to efficiently detect malicious activities by the nodes inside and outside the network [3]. As a matter of fact, the boundary of the network is not properly defined. Nodes can intermittently come into the network or leave it. Moreover malicious nodes can flood the network with junk packets hampering the network service or intentionally drop packets. But these nodes can but these nodes can subtly manipulate their harmful activities in such a manner that it becomes difficult to declare a node as malicious [4].

This paper elaborates the ongoing research on intrusion detection systems for detecting network layer attacks in mobile Ad-hoc networks. Precisely, GA and BFO protocol has been adopted and specifically monitors the vulnerabilities in the network layer. A solution is proposed based on the GA and BFO IDS technique for the detection of vulnerabilities in MANET.

## II. RELATED WORK

Dokurer et. al, (2007), investigated the effects of Black Hole attacks on the network performance. We simulated black hole attacks in Network Simulator 2 (ns-2) and measured the packet loss in the network. [8].

Anup Goyal and Chetan Kumar, (2010), has suggested a systematic learning method known as Genetic Algorithm (GA), to identify illegitimate nodes. The algorithm considers the varied features in network connectivity like protocol type, network service to destination and connection status to generate a type based rules. This was experimented by implementing in GA and trained it on the KDD Cup 99 data set to generate rules that can be applied to the IDS to categorize based on the attack types.[5]

K.S. Sujatha et.al (2012) propose a technique to analyze the exposure to attacks in AODV, specifically the most common network layer hazard, Black Hole attack and to develop a specification based Intrusion Detection System (IDS) using Genetic Algorithm approach. The proposed system is based on Genetic Algorithm, which analyzes the behaviors of every node and provides details about the attack. Genetic Algorithm Control (GAC) is a set of various rules based on the vital features of AODV such as Request Forwarding Rate, Reply Receive Rate and so on. The performance of MANET is analyzed based on GAC.[7]

Ahmed Shariff et. al, (2013), showed Mobile Ad-Hoc Networks (MANETs) are characterized by the lack of infrastructure, dynamic topology, and their use of the open wireless medium. Black-hole attack represents a major threat for such type of networks. The purpose of this paper is two folds. First, to present an extensive survey of the known black-hole detection and prevention approaches. Another objective is to present new dimensions for their classification.[6]

## III. DSR AND AODV

### 3.1 AODV

AODV is an on-demand routing protocol used in ad-hoc networks. This protocol is like any other on-demand routing protocol which facilitates a smooth adaptation to changes in the link conditions. In case when a link fails, messages are sent only to the affected nodes. With this information, it enables the affected nodes to invalidate all the routes through the failed link. AODV has low memory overhead, builds unicast routes from source to the destination and network utilization is less. There is least routing traffic in the network since routes are built on demand. When two nodes in an ad hoc network wish to establish a connection between each other, it will enable them to build multichip routes.

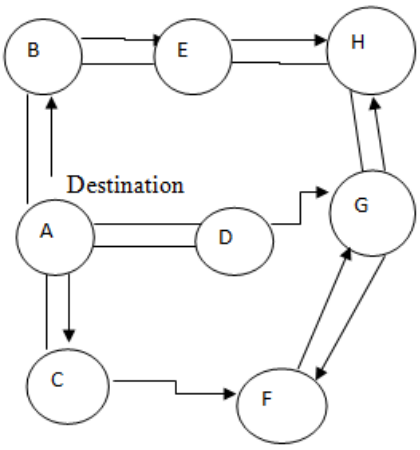



**Fig 1.5: AODV Protocol**

### 3.2 DSR

Dynamic Source Routing (DSR) is the very efficient protocol designed mainly for WSN, adhoc networks. Dynamic Source Routing (DSR allows network to be self- organizing as well as self- configured. The DSR contains two terms.

- Route Discovery
- Route Maintenance

Source routing helps to remove loops, packet forwarding etc. It is similar to AODV protocol in which demands are made on requirement. DSR also scales automatically, only when changes are needed. In DSR nodes forwards data packets from one node to another in order to enhance cooperation. AS sequence number is needed at destination, so rich topology is need like mesh, ring etc. Dynamic Source Routing (DSR protocol allows to search source node dynamically in whole network. Each data packet contains header that contains the information of destination as well as routing path. Thus there is no need to update regularly the routing table.

## IV. COMPARISON BETWEEN DSR AND AODV

Dynamic Source Routing (DSR) and Ad-Hoc on Demand Distance Vector Routing (AODV) are both routing protocols for wireless mesh/ad hoc networks. Both the protocols employ different mechanisms that result in varied performance levels. DSR and AODV can be compared and evaluated based on the packet delivery ratio, normalized MAC load, normalized routing load, and average end-to-end delay by altering the number of sources, speed, and pause time.

1. DSR has less routing overhead than AODV.
2. AODV has less normalized MAC overhead than DSR.
3. DSR is based on a source routing mechanism whereas AODV uses a combination of DSR and DSDV mechanisms.
4. AODV has better performance than DSR in higher-mobility scenarios.
5. DSR has less frequent route discovery processes than AODV.

## V. SIMULATION MODEL

A packet drop attack is also known as black hole attack in the network layer [9]. In black hole attack node drops packets at each step, then high loss of data takes place in the network. The node that drops the packet is malicious node. This attack can be viewed as following:

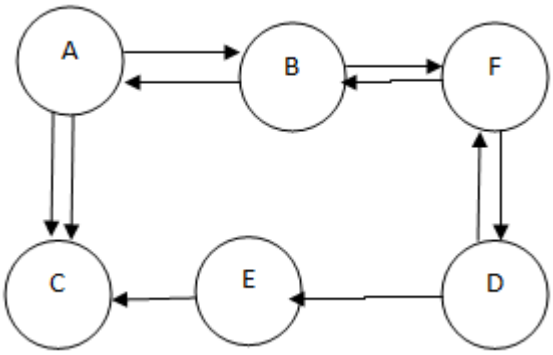



**Figure. 1 Black Hole Attack/ Packet Drop Attack**

Above figure shows that node A wants to send data to node D. So if node C is the shortest distance path from A to D, then it has to be followed. It will then receive the RREQ message from A node. As soon as node A starts to send the packet the node C drops packet in the middle of the data sending process [10].

Genetic Algorithm is a method of soft computing which uses the laws of selection and evolution. These algorithms are implemented by converting a problem in a particular field into a model a chromosome like structure. In computer network security, it is mainly used to find an optimal solution to a problem. The Genetic Algorithm starts by identifying a data set called population. Then these are individually encoded using bits, characters or integers and they form a chromosome. The next operation on them is an 'Evaluation Function' used to determine the genuine chromosome. During this process, two different operations namely, crossover and mutation are performed which is used to imitate the breeding and evolution. The selection of the chromosome is biased towards the fittest of the species. At last, the fit chromosome is selected once the optimization criterion is met. Below illustrates the basic functionality of Genetic Algorithm.

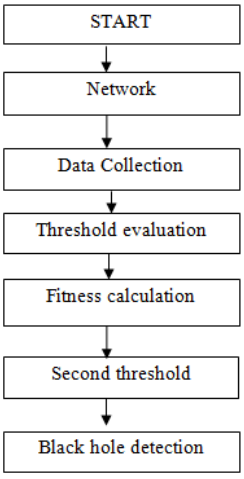



**Figure 2 Black Hole Attack/ Packet Drop Attack Proposed Flowchart**

The Figure elaborates on the functioning of Genetic Algorithm. The conditions for crossover if selected improperly will lead to the design of an illegal offspring. Hence identifying an optimal solution is not met if the finishing is slow or the convergence is premature. This is determined by means of selection of the fit condition to be either too large or too small.

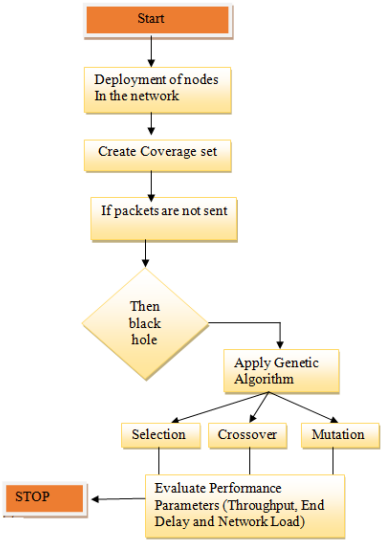



**Figure 3 Genetic Algorithms Processing For Black Hole Attack/ Packet Drop Attack**

## VI. RESULTS AND IMPLEMENTATION

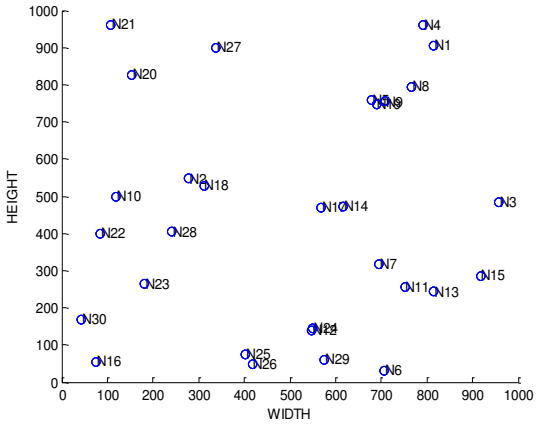



**Figure 1 AODV Network**

The above figure shows the AODV deployment of the nodes in the network. The area considered in 1000*1000 meters. The deployment of the nodes deals with the x locations and y location of the nodes.

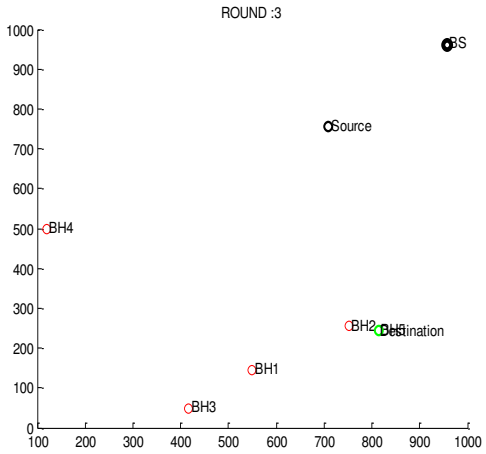



**Figure 2 Black Hole Nodes**

The above figure shows the black hole nodes at the end of the round because the nodes are the mobile nodes and their positions are changes according to the execution of rounds and the source in black color and destination in the green color is also plotted in the network

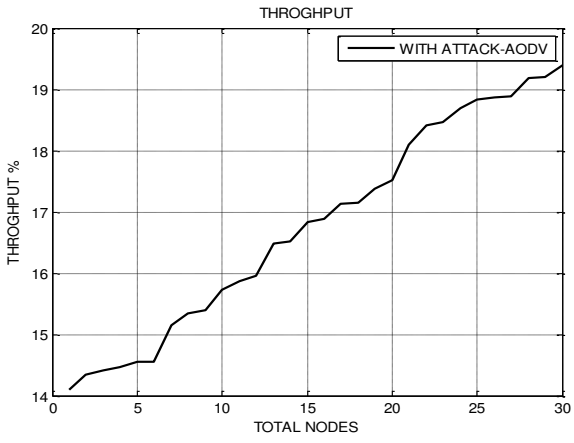



**Figure 3 Throughput AODV**

The above figure shows the Throughput with black hole attack using AODV and is having near about 19 percent which shows the effect of black hole attack in the overall performance of the network.

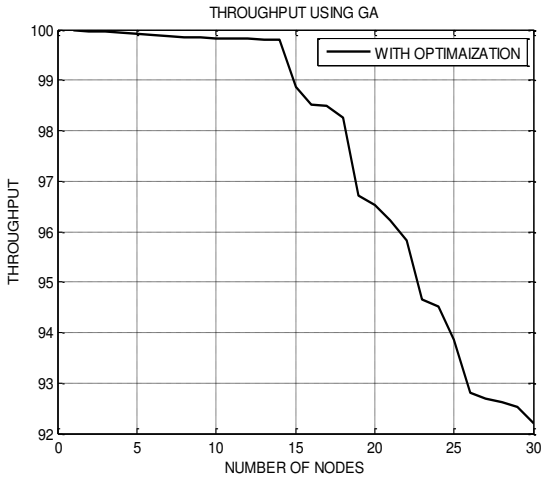



**Figure 4 Throughput Using Genetic Approach**

The above figure shows the throughput of the network using optimization approach using genetic algorithm which shows the overall performance of the network. This measure should be high for the efficient network. The graph shows the throughput 93.5% which is a sufficient measure to increase network lifetime.

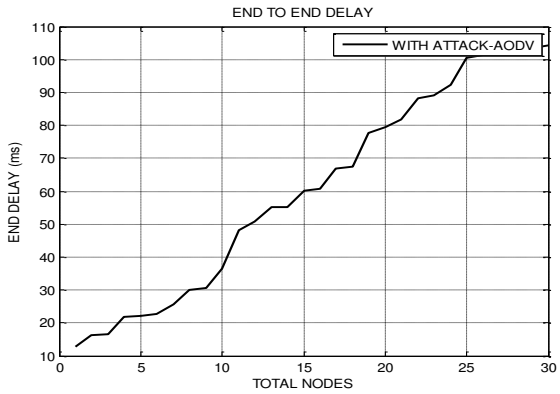



**Figure 5 End Delay AODV**

The above figure shows the End to end delay in milliseconds which is shows the successfully transmission with the particular time from source to the destination and shows 110 ms end delay to deliver packets from source to the destination

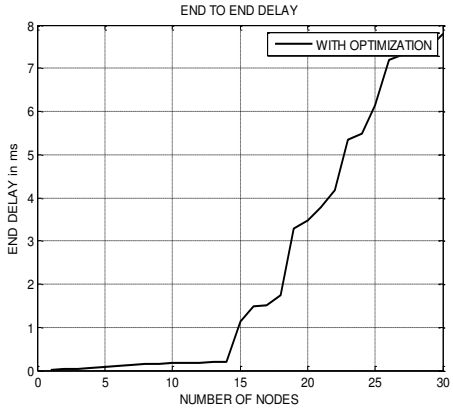



**Figure 6 Delay using Genetic Algorithm**

The above figure shows the end to end delay using genetic algorithm which is less as compared to the end delay with black hole attack. This measure should be low to successfully delivery of packets at particular time from source to the destination.

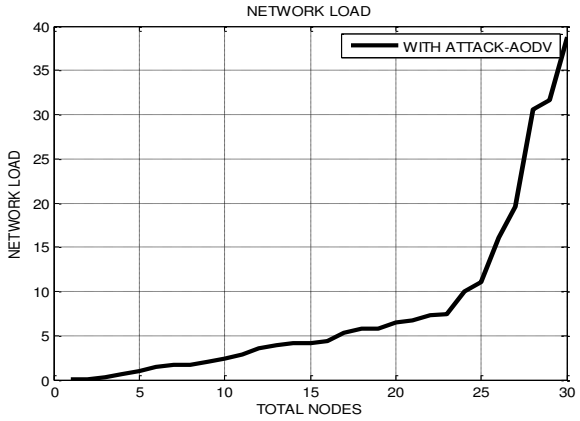



**Figure 7 Network load with black hole attack**

The above figure shows the black hole attack using AODV protocol in terms of network load which should be low for the efficient performance of the network

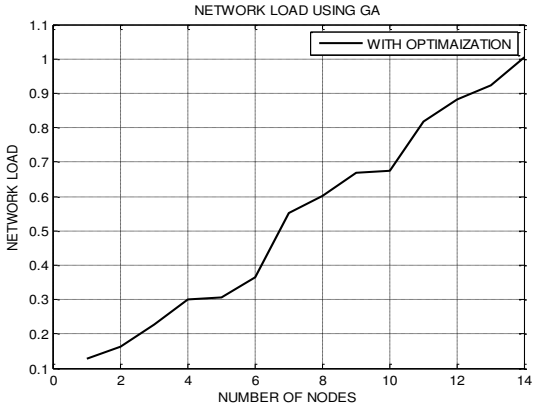



**Figure 8 Network Load Using Optimization Approach**

The above figure shows the network load using Genetic Algorithm which is used for the optimization and having less network load as compared to the AODV protocol in the presence of the black hole attack.

## VII. DSR RESULTS ANALYSIS

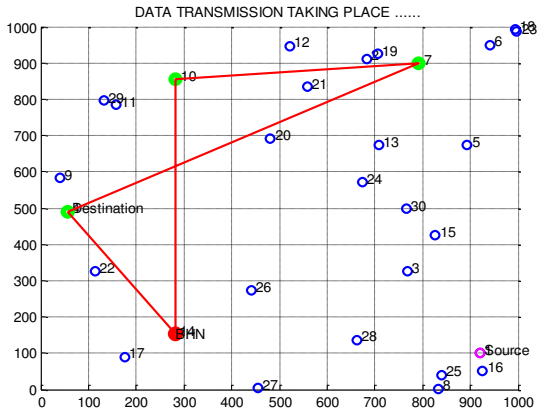


**Figure 9 DSR Network**

The above figure shows the DSR network architecture with deployment of the nodes in the network with source and destination plotted and transmission of the nodes are shown using black hole node in the network which is shows in red color in the above figure.

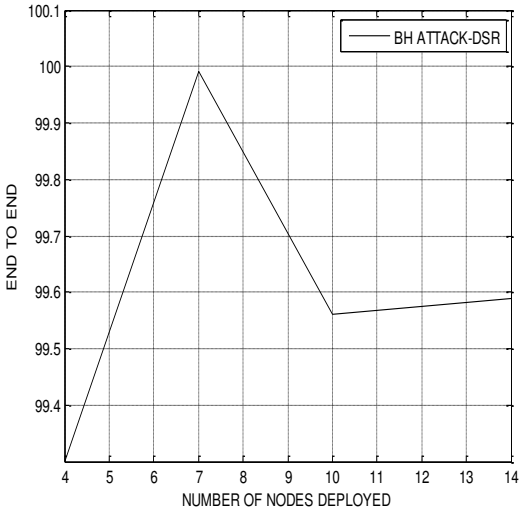



**Figure 10 DSR End Delay**

The above figure shows the End to end delay of Dynamic source routing in milliseconds which shows the successfully transmission within the particular time from source to the destination and shows 99.6 ms end delay to deliver packets from source to the destination

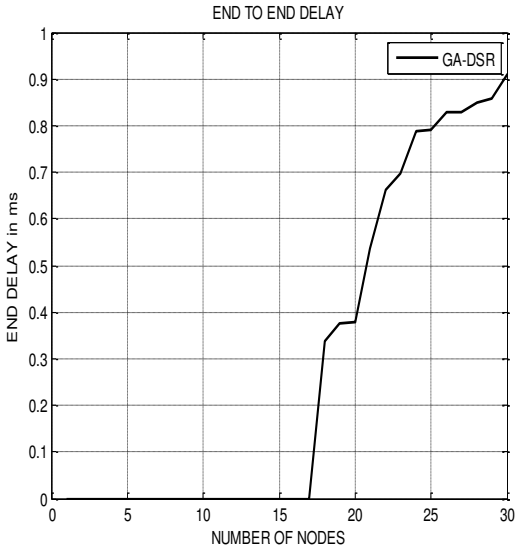



**Figure 11 End Delay with Optimization**

The above figure shows the end to end delay with optimization using Genetic algorithm and shows that the end delays less as compared to the DSR protocol in the presence of the black hole attack which is 0.9 ms.

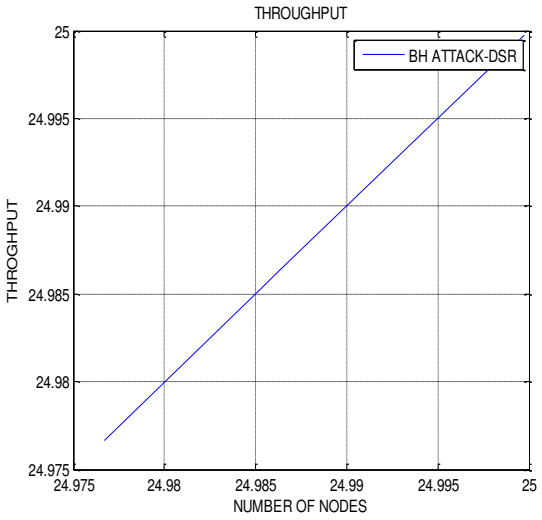



**Figure 12 Throughputs with Attack**

The above figure shows the throughput performance graph in the presence of the black hole nodes which is very less i.e. 25 % and should be high for the overall performance of the network

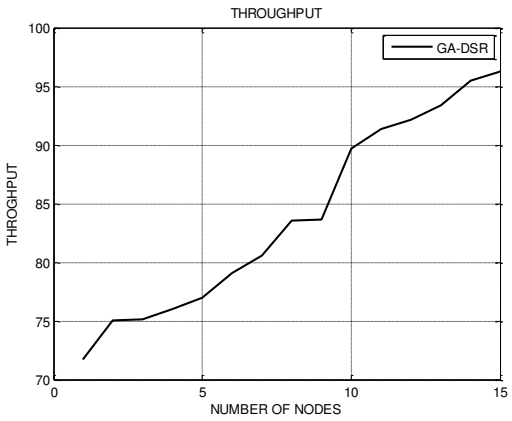



**Figure 13 Throughputs with Optimization**

The above figure shows the throughput performance graphs with genetic which is 97 % and is high to increase the network lifetime.

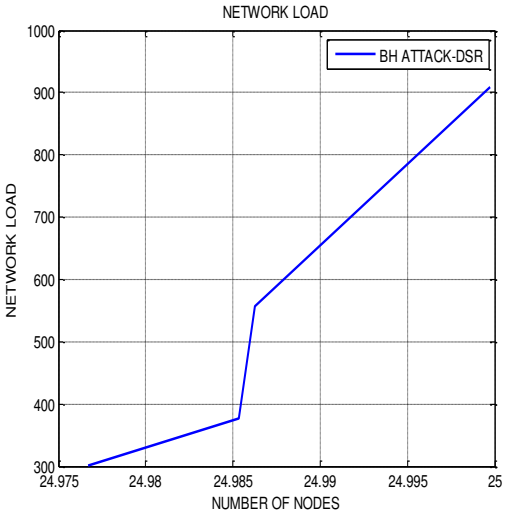



**Figure 14 DSR Network Load**

The above figure shows the network load in the presence of the black hole attack and is having very high value which degrades the performance of the network

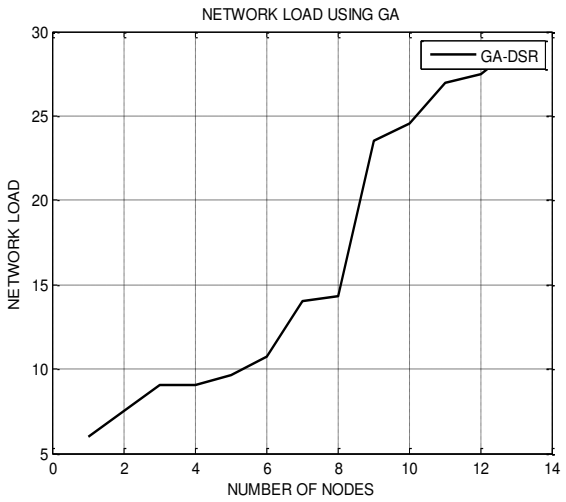



**Figure 15 DSR Network Load with GA**

The above figure shows the network load with genetic algorithm. Genetic algorithm optimizes the network load which is less as compared to the network load in the presence of the black hole attack.

## VIII. CONCLUSION AND FUTURE SCOPE

In this proposal, we have analyzed the effect of black hole attack in the performance of GA. The simulation has been done using the MATLAB. The simulation results show that when the black hole node exists in the network, it can be affected and decreased the performance of network and it can be optimized by using GA optimization algorithm. A hypothetical network was constructed for the simulation purpose and then monitored for a number of parameters. We simulate our model for various nodes. Initial position for the node is specified in a movement scenario file created for the simulation using a MATLAB. The nodes move randomly among the simulation area.So, the detection and prevention of black hole attack in the network exists as a challenging task. As future work, we intend to simulate and analyze the effect of the black hole attack in other routing protocols and we intend to perform the solution for the black hole attack and compare its performance with the AODV protocol and DSR protocol.

## REFERENCES

1. Sheenu Sharma and Roopam Gupta, “Simulation Study of Black hole Attack in Mobile Adhoc Networks”, In proceedings of Engineering Science and Technology. 2009.
2. Akansha Saini and Harish Kumar “Effect of Black hole attack on AODV Routing Protocol In MANET” , International Journal of Computer Technology, Volume1, Issue 2,December 2010.
3. Rajib Das, Dr.Bipul Syam Purkayastha and Dr.Pradipto Das “Security Measures for Black hole Attack in MANET: An Apporach” Internationsl Journal of Engineering Science and Technology, 2009.
4. P. Michiardi, R. Molva. "Simulation-based Analysis of Security Exposures in Mobile Ad Hoc Networks". European Wireless Conference, 2002.

5. Anup Goyal and Chetan Kumar, "GA-NIDS: A Genetic Algorithm based Intrusion Detection System", 2010.
6. Ahmed Sherif, Maha Elsabrouty. Amin Shoukry, "A Novel Taxonomy of Black-Hole Attack Detection Techniques in Mobile Ad-hoc Network (MANET)", IEEE, pp: 346-352, 2013.
7. K.S. Sujatha, V. Dharmar. R.S. Bhuvaneswaran, "Design of genetic algorithm based IDS for MANET", Conference: Recent Trends In Information Technology (ICRTIT), IEEE, pp.28-33, 2012.
8. Dokurer, S.; Erten Y.M., Acar. C.E., SoutheastCon Journal, "Performance analysis of ad-hoc networks under black hole attacks".Proceedings IEEE Volume, Issue, 22-25 March 2007 Page(s):148 –153.
9. Wei Li, "Using Genetic Algorithm for Network Intrusion Detection", 2010.
10. Ganapathy S, Yogesh P and Kannan A "Intelligent agent based intrusion detection system using enhanced multiclass SVM", Hindawi Publishing Corporation, Computational Intelligence and Neuroscience, 2012.
11. Revathi B, Geetha D, „A Survey of Cooperative Black and Gray hole Attack in MANET", International Journal of Computer Science and Management Research, Vol 1, no 2, September 2012.
12. Vijayan R, Mareeswari V and Ramakrishna K "Energy based trust solution for detecting selfish nodes in MANET using fuzzy logic", International Journal of Research and review in computer science, vol.2 No.3, June 2011.

# FABRICATION AND COMPARATIVE STUDY OF Zn-Ni METAL MATRIX COMPOSITE

**Manoj Kumar Pal[1], Arnav Vikram[2], Himanshu Pandey[3], Arka Ghosh[4]**

*[1]Assistant Professor, [2,3]Student, Department of Mechanical Engineering,*
*SRM University, NCR Campus, (India)*
*[4]Student, MBA, International School of Business & Research, (India)*

## ABSTRACT

*Metal Matrix Composites are in wide use these days. MMCs are light-weight, high strength materials that are increasingly being used for manufacturing in highly developed countries as they give products with better corrosion resistance, durability, portability, lightness, strength and are cost-effective too. This paper presents study on the Zn-Ni metal matrix composites. In this paper the properties of Zn-Ni metal matrix composite are reviewed. The fabrication of Zn-Ni metal matrix composite was done and their properties were studied and comparisons were drawn. Four specimens with different Zn-Ni compositions 10%, 20%, 30%, 40% of nickel were made and were tested for their properties. Each specimen showed different properties and appearance. The results indicate that the method chosen was successful and how the properties varied on changing Ni compositions. Four tests were performed on each of the specimen and graphs were plotted for each property showing the variation in property as Ni percentage gets changed. The tests performed were Rockwell hardness test, Brinell hardness Test, Ultimate Tensile Strength and Thermal conductivity test. The readings of each test were noted and analyzed. After looking and investigating all these results we inferred that the specimen Zn-Ni MMC used in various industries across the world should be increased especially in developing countries like India where cost-effectiveness is a big factor in industries vide percentage of Ni that can be chosen as per the application.*

***Keywords: Metal Matrix Composite, Zinc (Zn), Nickel (Ni), Rockwell Hardness, Brinell Hardness***

## I. INTRODUCTİON

Metal Matrix Composite (MMC) is a combination of two or more metal. It gives the unique combination of properties. Composites are in use since very long time in industries. Today MMCs are used in the industrial applications spanning aerospace, sports, automobiles etc. The main aim of making a composite is for properties like harden-ability, ductility, light-weight and wear-resistance. In composite materials, the reinforcement can be particle, whisker or fibres. The matrix metal can be ceramic, plastic or metal. The reinforcement can be made from ceramic, plastic or metals. Today, lot of research work is going on in the discovery of metal matrix composite globally. Metal matrix composites are made of continuous metallic matrix and one or more discontinuous reinforcing phase. Researchers are giving attention to the production of MMCs at an accepted level and at a low cost. Composites give good properties of material like high strength, high modulus, high toughness, thermal shock and good thermal conductivity. There are different types of metal matrix composites on the basis of reinforcement viz. particle reinforcement, fibre reinforcement, continuous reinforcement, and

laminate reinforcement. Various process are used for making the metal matrix composite like liquid fabrication method, solid state fabrication method, metal injection mouldings, friction stir process, mechanical alloying, etc. There are a lot of techniques for fabricating MMC like stir casting, powder metallurgy, gas infiltration, spray forming, chemical vapor deposition, electroplating, etc.

Out of all these techniques, stir casting is the most common as it is simple to use as well as it is cost effective in nature. The advantage of this process is simplicity, flexibility and applicability to a large quantity of production. It is very attractive because of minimized final cost of product. There are many factors that need considerable attention, which primarily includes achieving a uniform distribution of reinforcement. During these conventional processes, the major difficulties faced are improper wetting due to oxidation which exhibit interface binding between matrix and ceramic phases. Improved wetting must be achieved to obtain a good bond between matrix and reinforcement.

Composites are used in automobiles like engine of car and aerospace applications. Zinc is very shiny bluish white colored material and it has ability to resist corrosion and protect the other metal. As Zn is abundant in nature, the composite Zn-Ni is used in various applications. For reinforcement, Ni is very good metal as it is very hard and ductile in nature. This metal gives optimum reinforcement results. It is a shiny white lustrous metal. For making MMC, reinforcement metal should be taken in the form of powder because Ni is very reactive and easily oxidizes in presence of $O_2$. It can be used in protection from rusting. Hence, this composite can also be used for alloys, batteries, electroplating process and casting process. In this paper, we have discussed about fabrication and characterization of MMC (70%Zn+30%Ni) and its applications. Composite materials have been used to solve technological problems since a long time but it gained popularity back in 1960 and till date continues to intrigue researchers.

## II. FABRICATION

There are a lot of methods for fabricating metal matrix composites viz. Liquid phase process, Solid- solid process, Deposition process, In-situ process, Solid – liquid process.Out of all these processes, we have chosen liquid phase process. This process consists of a method known as stir casting. In this process a dispersed phase and a molten metal matrix are mixed together with the help of mechanical stirring. In our case, molten matrix was Zn and dispersed phase was Ni. The dispersed phase was in form of particles and this particulate dispersed phase was chosen as they get mixed with the matrix easily and can provide uniform distribution of the reinforcing material which in turn would provide constant mechanical properties of material at all points. In stir casting process (see fig. 1), the Zn was melted in the furnace at a temperature of 700°C in a graphite crucible and then pre-heated Ni particles were poured in the same crucible containing the molten Zn and temperature was increased till 1000°C. The amount of nickel was put one by one in percentages of 10%, 20%, 30% and 40%. After that the mixture of both the metals was stirred at rpm ranging from 200-1000 for some time. When complete mixture of zinc and nickel was done properly then specimen with 10%, 20%, 30% and 40% Ni were sand casted and 4 different specimens of composites were prepared (see fig. 2).

The relation between the properties and cost of the constituent material of the MMC was determined by the fabrication process used. Now powder metallurgy is a form of solid-state fabrication process of composite in which mechanical mixing of matrix phase powder and reinforcement phase powder takes place. Then the consolidation of the powder mixture rakes place with the help of hot pressing which normally takes place above

the solidus temperature or extrusion is applied. Now basically there are two techniques of incorporating the reinforcement in the matrix viz. a) Ex-situ process- In which incorporation of reinforcement particle is done with the help of stirring, blending or injection and b) In situ process- here the reinforcement phase is prepared by some reaction process between constitutes of MMC to form carbides, nitrides, silicates, borides. The ex-situ process has a lot of problems like micro porosity, uneven distribution of reinforcing material, cost of processing etc. while the in-situ process has several advantages like improved fatigue and creep resistance, high interfacial strength, good wet ability, corrosion material, lower cost of production. But this advantage of in-situ process may not be applicable for all system of composites as it is system and property specific.

**Figure 1. Stir casting process.**

**Figure 2. Zn-Ni composite with different Ni %.**

## III. TESTS & ANALYSES

The following experiments are conducted to study the properties of different percentages of Zn-Ni metal matrix composites. Tests done on those 4 specimens were Tensile Testing, Rockwell Hardness Testing, Brinell Hardness Testing, Thermal Conductivity Testing and Micro Structural analysis.

### 3.1 Tensile Testing

A tensile test measures the resistance of a material to a static or slowly applied force. A machined specimen is placed on the testing machine and load is applied. All the specimens that are 90%Zn+10%Ni, 80%Zn+20%Ni, 70%Zn+30%Ni, 60%Zn+40%Ni are one by one tested on the Universal Testing Machine. A strain gage or extensometer is used to measure elongation. The stress obtained at the highest applied force is the Tensile Strength. The Yield Strength is the stress at which a prescribed amount of plastic deformation (commonly 0.2%) is produced. Elongation describes the extent to which the specimen is stretched before fracture. Information concerning the strength, stiffness, and ductility of a material can be obtained from a tensile test. The readings and the corresponding graph we got are in Table 1 and Graph 1 respectively.

### 3.2 Rockwell Hardness Testing

The Rockwell Hardness test makes use of indentation to measure the hardness of the different specimens. This test performs with different load. Firstly specimen is placed on the indenter and minor load (10Kg) is applied on the specimen and release with major load. This test measures the permanent depth of indentation produced by a force on an indenter. In Rockwell testing machine there are three different scale reading for load like 100Kg, 150kg, 60Kg etc. Generally scale c is used to measure the readings. This test is easier to perform and more accurate and other type of hardness testing methods. For each specimen it shows different scale readings. We

noted down different readings for each specimen. It measures the hardness of specimen and gives the accurate results. The readings and the corresponding graph we got are in Table 1 and Graph 4 respectively.

### 3.3 Brinell Hardness Testing

Brinell Hardness Test is more or less same like Rockwell Hardness test but the difference is that we use a formula to get the hardness in units such as $kg/mm^2$. Steel ball indenter was used here and a load of 3000kg was used on the four composites specimens. The diameter of the ball was 10mm. Then readings were found out for all the specimens and noted in Table 3. This test is used to determine the Brinell hardness. In this test specimens are loaded on indenter with minor load and released with major load. This test offers very high load 3000kgf and a 10mm wide indenter. This test applies a predetermined load (f) to a carbide ball of fixed diameter (D) which is held for a predetermined time period and then removed. Chart is used to convert average diameter measurement to a Brinell hardness number. In this test, force applied ranges from 500 to 3000kgf. This test measures the permanent width of indentation produced by a carbide indenter applied to test specimen at a given load, at given time period. An indentation is made from Brinell hardness testing machine and then measured for indentation diameter in second step. The resulting measurement is converted to a Brinell value using the Brinell formula or conversion based chart formula. This test will use 3000kg load with 10mm ball diameter. Brinell test ranges from 3000kgf-1kgf and ball indenter diameter ranges from 10mm-1mm and is used if it is known that the shorter time period does not affect the result. The test metrics are ball diameter (D), impression diameter (d), load (l) and Brinell Hardness (HB). The readings and the corresponding graph we got are in Table 1 and Graph 2 respectively.

### 3.4 Thermal Conductivity Test

This test tells us about the extent of a material conducting heat. The material with known conductivity is placed on both sides of fabricated specimens. The thermal conductivity of the material depends upon the thermal properties and intermediate temperature. The conductivity of material may vary with change in temperature. In this test, thermal conductivity slowly increases and then decreases. This test shows conductivity of each material varying with change in properties and temperature. The test is generally of two type's viz. Steady State Method and Non-Steady State Method. The heat transfer takes place with the help of heater and process of heating was convection. When a specimen has a steady state then readings are noted down in table 4and a graph is plotted too. The readings and the corresponding graph we got are in Table 1 and Graph 3 respectively.

### 3.5 Microstructure Analysis

The microstructure of the specimens Zn-Ni 10% Zn-Ni20% Zn-Ni30% Zn-Ni40% were made in the lab. The microstructures cannot be seen by naked eye. The microstructure was obtained through belt polishing, sand paper, double disc polishing, and etching (serially). Polishing methods are done to remove abrasiveness of specimen. Grain boundaries are visible on placing the specimens under microscope which is not possible elsewise. Under the microscope, we see the color of specimens and their fine surfaces. First we check in our microscopic study there should be no grain boundary inside the specimen and then we do the study. The processes involved in our study are:-

**Belt polishing:** In this method, specimen is polished till the surface of specimen gets cleared of abrasives. While doing this process, water is added when specimen gets heated. Adding of water also cleanses the surface of the specimen and cools the specimen.

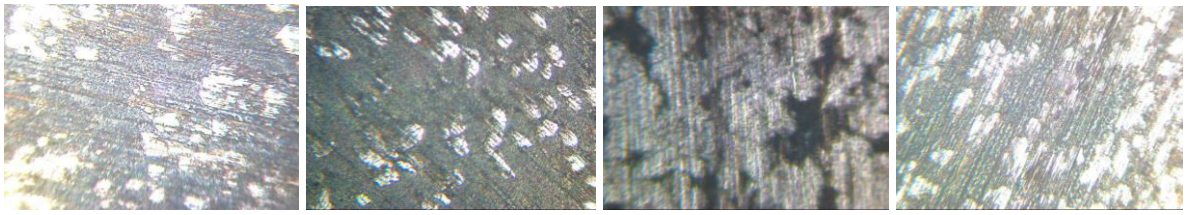
**Hand Polishing:** It is done for polishing the specimen. This process is done after belt polishing and sand paper is used for this purpose. First specimen is polished on abrasive paper and then polished on different grades of sand paper.

**Double-Disc Polishing:** The polishing of specimen is done till scratches on specimen surface are eliminated completely. Water is sprinkled on the specimen to get smooth surface and for cooling.

**Etching:** This process is done at the end. This process involves sprinkling of Keller's reagent over the specimens. The time for etching is generally between 15-30 seconds. After etching is done, the specimen is washed in water and then it is dried with the help of a blower. Finally the specimens are ready for studying their microstructures under the microscope.

Different specimen shows different microstructure caused due to the different Ni %. Under microscope we see the color of specimen and grain boundaries. In the microscopic study, we found that 30% Ni gives greater tensile strength and lower thermal conductivity. Through microscope we saw that all the specimens show original color and fine structure. It shows the inherent properties of specimen.

We see that all four specimens show different microstructures (ref. Fig. 5). The variations in the microstructure of the specimens are due to different Ni % incorporated during fabrication time. Different specimens show different microstructure. On analyzing this test we saw that the greater Ni % shows clear and fine picture of microstructure. In the microscope it shows better result than other specimen. In other specimens we can see the Ni but it's not better than Ni 40%. Nickel is a silver lustrous metal with golden tinge. Under the microscope, we can clearly see the fine and clear picture of the specimen. There is no grain boundary in the specimen. We observed that the property of each material is different and each give different results. According to our microscopic analysis, we found that microstructure of nickel can easily be obtained and it has a golden silver color. Thus, our specimen gave microstructure visible through microscope and at different Ni % easily saw the grains of structure.



**(a)** 90%Zn-10%Ni **(b)** 80%Zn-20%Ni **(c)** 70%Zn-30%Ni **(d)** 60%Zn-40%Ni

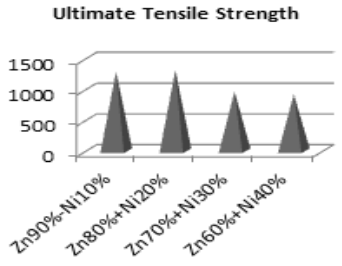
**Figure 3. Micro Structural View of all the Four Fabricated Specimens**

## IV. TABLES & GRAPHS

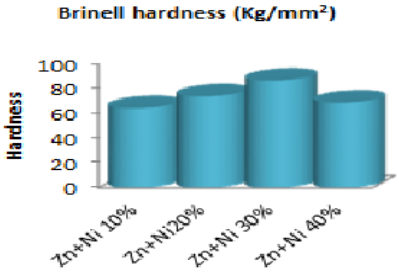
**Table 1. Readings of the Tests Corresponding to Various Ni %**

| Ni % | Ultimate Tensile Strength (N/mm²) | Rockwell Hardness No. (HRC) | Brinell Hardness (kg/mm²) | Thermal Conductivity (W/mK) |
|---|---|---|---|---|
| 10 | 1290 | 45 | 64.56 | 70.45 |
| 20 | 1312 | 41 | 74.12 | 70.85 |

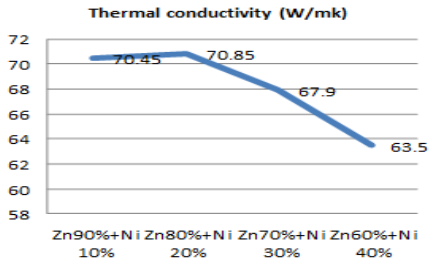
| | | | | |
|---|---|---|---|---|
| 30 | 965 | 32 | 87.05 | 67.90 |
| 40 | 920 | 28 | 69.15 | 63.50 |



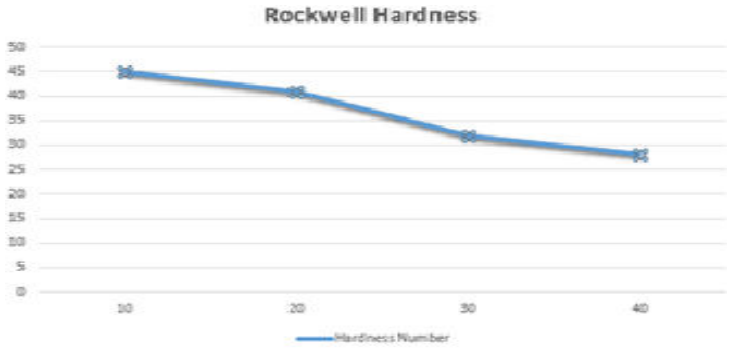

Graph 1




Graph 2




Graph 3




Graph 4

## V. RESULTS

After careful examination and testing of all the fabricated MMCs, we can say that due to the different percentage of Nickel in different specimens there are variations in their properties. This behavior can be attributed to the interaction between Zn and Ni at different compositions. These variations in the properties of this metal matrix composite are also due to the parameters like reinforcement type, manufacturing process type, defects, Ni %, particle size and impurities. Properties of different specimen were noted down and were plotted on a graph. There was inconsistency in the graph. So, to get an idea of the testing values in a nutshell a comparison chart was plotted (see fig. 4). As observed from the obtained results, Rockwell hardness decreased, Brinell hardness increased and then slumped, thermal conductivity and tensile strength decreased with increase in Ni %. On observing keenly, we found that the thermal conductivity decreases with increase in % of Ni as is depicted by the curve of thermal conductivity. Hardness is best at 20% Ni. Thus, depending on the application, composition can be chosen accordingly.

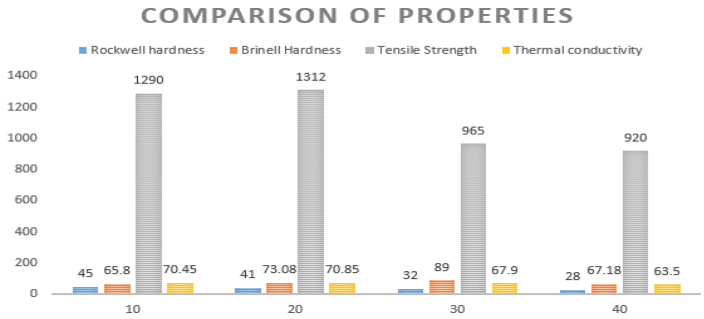



**Figure .4 Comparison Chart**

## IV. CONCLUSION

The specimens Zn+10% Ni, Zn+20%, Zn+30% Ni, Zn+40% Ni were made. After testing of the specimens with different percentages of the nickel that are 10%,20%,30%,40%, results were noted down and were thoroughly analyzed. We found that specimen properties like Rockwell hardness, Brinell hardness, tensile strength thermal conductivity varied with variation in Ni % and hence usability will vary according to applications and type of industry. But, after looking and investigating all these results we can say that the specimen Zn-Ni MMC use in industries should be increased especially in developing country like India where cost-effectiveness is a big factor in industries as this MMC has wide and varied applications in automotive industries as cylinder walls, engine cooling packs, cycle frames, pistons and in other industries too.

## V. ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present this research work. We are grateful to all the faculties of Department of Mechanical Engineering, SRM University (NCR Campus) for the constant support and guidance throughout the course of our work. Their sincerity, thoroughness and perseverance have been a constant source of inspiration for us. A special and heartfelt gratitude to Mr. Manoj Kumar Pal, who is also the co-author of this paper. It is through his cognizant efforts that our endeavor has seen the light of day.

## REFERENCES

[1] Reddy, S.U., Synthesis, mechanical and tribological characteristics of Mg-SiC metal-matrix composites, M.Tech Thesis, Department of Mechanical Engineering, National University of Singapore, 2005.

[2] Telang A.K., Rehman A, Dixit G, Das S, "Alternate materials in automobile brake disc applications with emphasis on Al composites: A technical review", Journal of Engineering Research and Studies 2010, 1 (1), 35–46.

[3] Singla M, Dwivedi D.D, Lakhvir S, Chawla V, "Development of aluminium based silicon carbide particulate metal matrix composite", Journal of Minerals and Materials Characterization and Engineering, 2009, 8 (6), 455–467.

[4] Tan, M.; Xin, Q.; Li, Z.; Zong, B.Y., "Influence of SiC and $Al_2O_3$ particulate reinforcements and heat treatments on mechanical properties and damage evolution of Al-2618 metal matrix composites", Journal of Materials Science 2001, 36, 2045–2053.

[5] Shorowordi K.M, Laoui, T, Haseeb A.S.M.A, Celis J.P, Froyen L, "Microstructure and interface characteristics of $B_4C$, SiC and $Al_2O_3$ reinforced Al matrix composites: A comparative study", Journal of Materials Processing Technology, 2003, 142, 738–743.

[6] Lloyd D.J, Lagace H, McLeod A, Morris P.L, "Microstructural aspects of aluminium-silicon carbide particulate composites produced by a casting method", Materials Science and Engineering, 1989, 107 (A), 73–80.

[7] Aqida S. N, Ghazal M. I, Hashim J, "Effects of porosity on mechanical properties of metal matrix composite: An overview'', Jurnal Teknologi A, 2004, 40 (A), 17–32.

[8] D. B. Miracle, "Metal Matrix Composites from Science to Technological Significance," Composites Science and Technology, Vol. 65, No. 15-16, 2005, 2526-2540.

[9] J. Hashim, L. Looney, M. S. J. Hashmi, "Metal Matrix Composites: Production by the Stir Casting Method," Journal of Materials Processing Technology, Vol. 92-93, 1999, 1-7.

# REVERSE LOGISTICS FOR PLASTIC REPROCESSING INDUSTRIES IN SOUTH INDIA

**S. Robin Divahar[1], C. Sudhahar[2], Karthikeyan.L.M[3]**

[1] *Research Scholar, Anna University, Chennai, (India)*

[2]*Department of Mechanical Engineering,*

*Marthandam College of Engineering and Technology, (India)*

[3]*Department of Aeronautical Engineering, Techno Global University, (India)*

## ABSTRACT

*In South Asian countries like India, small plastic containers are mostly used for packing the consumer products like food, medicine and chemicals. The chemical and waste storage containers are mostly manufactures from recycling plastic granules. In India, recycling plastic granules are obtained from used plastic products and they are collects through reverse logistic process. After an extensive literature survey, the possibilities of complete recycling and remanufacturing of plastic containers in the market by reverse logistics has analyzed and a new logistical method has proposed.*

***Keywords: Logistics, Reprocessing, Plastics***

## I. INTRODUCTION

In South Asian countries like India, recycling plastic granules are obtained from used plastic containers and they are collects through reverse logistic process. After an extensive literature review, this study aims to understand the importance of plastic reverse logistic process from recycle plastic manufacturers in Kanyakumari. It is a small district cited in the southern tip of Indian cape of South Asia. This region covers 1671 square kilo meters with 1,870,374 numbers of people as per 2011 census. Nagercoil and Marthandam are the major town areas in this district. In Kanyakumari, most of the plastic industries are small scale industries (SSI) or tiny units and the manufacturers are making the plastic containers and bottles for non food processing goods. The one time use products like waste chemicals and bath room cleaning powders need this type of plastic bottles. Also this types of products need not to pack in high quality packaging as per the quality and economical requirements. After the completion of the technical and marketing survey, different types reverse logistic methods has identified and studied for plastic recycling industries in Kanyakumari to collect the used plastics effectively for recycling. The identified process has useful for remanufacturing of returned goods for manufacturers in economical and quality point of view.

## II. PLASTIC MANUFACTURING INDUSTRIES IN SOUTH INDIA – A BRIEF STUDY

The major types of plastic moulding process are blow moulding, injection moulding and extrusion moulding. The small scale industries involving into the above moulding process are always have to concentrate in purchase, sales, manufacturing, marketing and transporting the products as shown in the figure 1. The major

duties of purchase department involves the purchase of processing machineries, raw materials and accessaries for production purpose.

The sales department always concentrates to sale the manufacturing goods for gain the ultimate profit with the desired quality. Each type of plastic products needs seperate moulds or dies to obtain the final shape. The plastic manufacturers may have seperate mould making units or they will create the moulds from seperate mould makers. The blow, injection or extrude machineris involves the manufacturing duties for procuct making. The quality control and coustomer care jobs have controled by marketing departments. Forward and reverse logistic assignments can handled by logistics departments.

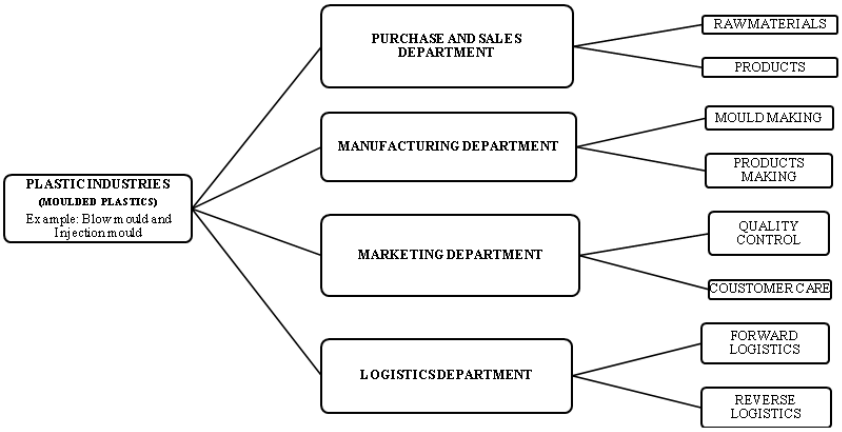



**Figure 1. Various Departments in Plastic Processing Industries**

In Kanyakumari district, more than 45 plastic and petro based small and micro scale industrial units are available. Nearly 800 people are getting employment from these units. In this district, nearly 40% of the plastic manufacturers are manufacturing plastic products only from reprocessing plastics for non food packaging or consumable applications because of the following reasons.

- Plastic consumers are expecting low price but high quality products for domestic and industrial applications. But the products of above quality have to make from fully automatic machineries with ultimate production capacity and well equipped advance production systems. These types of advance systems are available only in major cities and so that most of the high quality products are delivered for local consumers only from major cities.
- Because of the well planned logistical systems and reduction in product cost of imported items, local consumers preferring imported and transported plastic items for non packaging applications.
- In kanyakumari district, most of the plastic moulding products are manufacturing only by semi automatic or manually operated machineries suitable for packaging applications. The local manufacturers are fully depended on repackaging traders and small scale industrial manufacturers around the township.

## III. LOGISTICS PLANS FOR PLASTIC MANUFACTURING INDUSTRIES

Henrique Luiz Corrêa and Lucia Helena Xavier [1] have mentioned the concepts, design and implementation of reverse logistics systems for sustainable supply chains in Brazil. In South India, the logistic plan of SSI plastic moulders has always prefers to reduce the transportation cost. Patricia Oom do Valle, Joao Menezes, Elizabeth Reis and Efigenio Rebelo [2] have studied the customer involvement in the reverse logistics system for recycling household packaging. As per the common logistical method used in South India (known as ‘method

1'), the supplier will choose the way of transportation to distribute the goods by covering maximum customers in a single transportation in a specified area. In this method the return products will carry by the supplier with separate logistics providers. Marisa P. de Brito, Simme D. P. Flapper and Rommert Dekker [3] indicated the reasons for product returning. They are manufacturing returns, commercial returns, product recalls, warranty returns, service returns, end of use returns and end of life returns. Leo Kroon and Gaby Vrijens [4] have studied the reverse logistics of returnable containers. Aarón D. Castillejo T and Fredrik Stensson [5] have indicated the returnable plastic packaging flow in the automotive industry. In another logistical method, (known as 'method 2'), the delivery as well as return the products will carry in same vehicle for consumers and suppliers. To minimize the transport cost, the plastic distributors always prefers this logistic plan known as single way logistic system (SWLG) by carrying the return good after the product delivery. The description of SWLS has shown in figure 2.

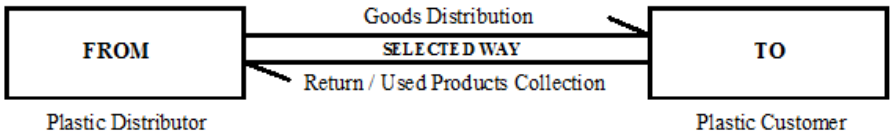



**Figure 2. Basic Concept of SWLG**

Reza A. Maleki and Jonathan Reimche [6] have studied the logistic and returnable container's physical flow through flow diagram. The blow moulding plastic goods like plastic containers has less weight but covers more space in the stuffing area of vehicle. It creates high transport cost during the time of the delivery of plastic goods. For this case, to control the logistic cost, the following factors must consider during the single way transportation.

- Ensure full load in vehicle for various customers in the same route
- Effective distribution for all customers on the way in simple trip
- Effective return trip with the return goods from the customers.

SWLS covers the forward and reverse logistics of goods in a single trip. The forward trip of SWLS is for deliver the manufacturing goods to multiple buyers and the reverse trip is for collect the used or returned goods. The major consideration for SWLS for plastic industries as follows.

- Trip for delivery goods must arrange only with full stuff in the vehicle
- The trip must covers only selected areas with effective delivery and collection
- The charge of return vehicle must reduce by the effective collection of return goods

Mikko Kärkkäinen, Timo Ala-Risku and Marianna Herold [7] have studied for managing the rotation of reusable transport packaging. Figure 3 shows the step by step approach for single way logistics with multiple customers.

## IV. REVERSE LOGISTICS (RL) METHODS FOR REPROCESSING PLASTIC GOODS

Isabel Fernández Quesada [8] has reviewed various papers for study the concepts of reverse logistics. Dale S. Rogers and Ronald S. Tibben-Lembke [9] has studied the reverse logistic activities, reverse flow of goods, classification of reverse logistics activities and importance of reverse logistics. Magdalena Graczyk and Krzysztof Witkowski [10] have indicated the value chain and different kinds of recovery in integral supply chain for plastic products. The production of plastic goods like non food packaging containers needs

reprocessing plastic raw materials (RP materials). So that the recycle plastic product manufacturers have separate RP materials making unit with the other departments indicated in figure 1. Brandon Kuczenski and Roland Geyer [11] have indicated the flow diagram for pet bottle reverse logistic program. The RP materials can obtain from used plastic products through the following five steps involved in recycling process. They are used plastic collection, manual sorting, chipping, washing and pelleting. Chee Wong [12] explains the plastic recycling supply chains and waste handling. The one time used plastic products can obtain from used plastics collection shops. The reprocessing plastic makers have second quality plastic production machineries for produce the raw materials from used plastic products. Yukie Umeda, Hiroshi Tsukaguchi and Yan Li [13] have mentioned the various methods for efficient collection and reverse logistics system of electrical appliances recycling. For used plastic collection process, the manufacturer or logistics provider collects the used plastic from collection shops or from the consumers. This collection process also can carry in the same trip with the product delivery and returned product collection of 'method 2'. This is the third logical method and in this method, the trip will be arranged periodically as per the following requirements

- Customer's product requirements and availability
- Used and returned products availability with customers and / or in collection shops

The comparison of the above three logistical methods as mentioned below.

**Method 1:** Sending the products to the consumers

**Method 2:** Sending the products from manufacturers and collection of returned products from consumers in a single trip

**Method 3:** Sending and collection of products from manufacturers, consumers and waste collection shops in a single trip (suitable for reprocessing plastic industries)

**Table 1: Transport Frequency with Respect to Distance**

| Sl.No | Distance (kilometres) | Vehicle | Operation frequency |
|---|---|---|---|
| 1 | 0 < 5 | Light vehicle | frequently |
| 2 | 5 to 10 | Medium vehicle | Weekly or monthly |
| 3 | > 10 | Heavy vehicle | Well defined time interval |

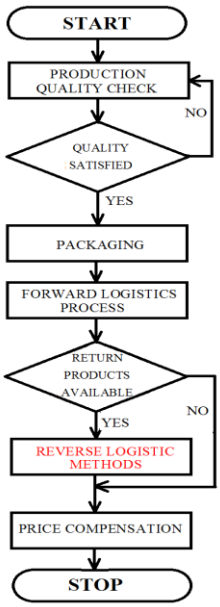



**Figure 3. Step by Step Approach for Single way Logistics with Multiple Customers**

## V. ROUTE MAPPING

Georgiadis.P and Vlachos.D [14] has done the numerical investigation for closed loop supply chain modelling for a reverse logistic dynamic system. Before deliver or return the goods, the route arrangement must define by the transport provider for effective and reliable logistic. Dennis W. Krumwiede and Chwen Sheu [15] mentioned the reverse logistics decision-making model for logistics entry by third-party providers. Most of the SSI plastic manufacturers in Kanyakumari district have their customers around 15 to 20 kilo meters from their production units as shown in figure 4. A typical unit as shown in figure has 11 consumer units and 6 return products collection areas. In this example, the production industry has 4 consumers and 2 recollection units in north-eastern side, 3 consumers and 1 recollection unit in north-western side, 2 consumers and 2 recollection units in south-eastern side and also have 2 consumers and 1 recollection unit in south-western side.

**Table 2: Trip Arrangement with Respect to Direction**

| Sl.No | Trip | Region | Direction | Number of Consumers covered | Number of Collection units covered |
|---|---|---|---|---|---|
| 1 | A | 1 | NW | 3 | 1 |
| 2 | B | 2 | NE | 4 | 2 |
| 3 | C | 3 | SW | 2 | 1 |
| 4 | D | 4 | SE | 2 | 2 |

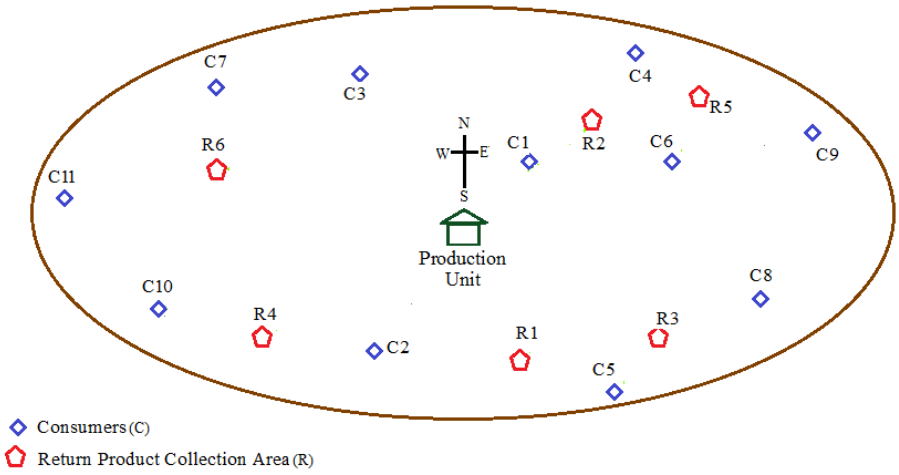



**Figure 4 Sample Delivery and Collection Areas**

Figure 5 indicates the delivery or collection units as per distance measurements. In the measured unit, 2 consumers and 1 recollection units are available within 5 kilometres. The manufacturer can use the small size transport vehicles for the delivery or collection duties for these units frequently. But for the units available more than 10 kilometres, the heavy vehicles are suitable in the well define time interval to save the transport expenses and to achieve the effective logistic solution. The different types of suitable vehicles with respect to the distance covers by destination as mentioned in table 1.

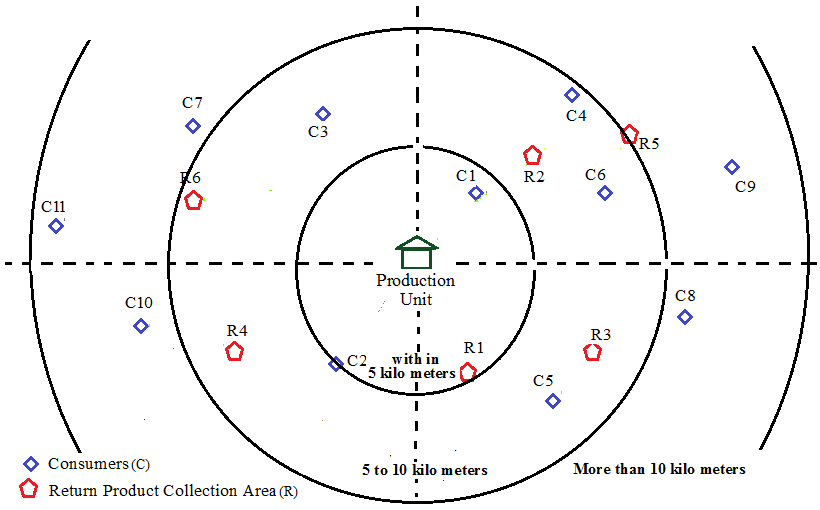



**Figure 5. Mapping with Respect to Distance**

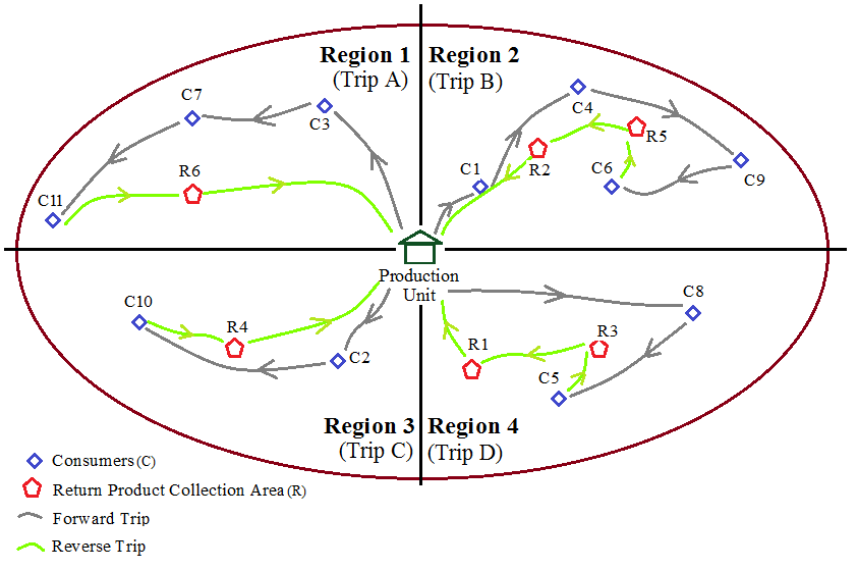



**Figure 6. Route Mapping with Respect to Transport Direction**

The distance mapping may not suitable for SSI industries unless they have more number of consumers in the above three regions. If the production or collection orders are below the production capacity, the mapping can do with respect to the direction of trip. The figure 6 indicates this type of mapping covers the customers with respect to north-west (NW), north-east (NE), south-west (SW) and south-east (SE) directions. The details of trip has mentioned in the table 2. At least one collection units must cover in each trip and the ultimate prospect of this arrangement is the full stuff in forward and reverse transport with delivery and collection of goods.

## VI. CONCLUSION

The proper scheduling in reverse logistic for reprocessing SSI plastic manufacturers is the most important process. This study indicates the various plans for collect and transfers the returned goods during the time of product delivery. The effective transport plans always importance for the growth of industries. The typical plans indicate in this paper explained as the examples. The manufacturers can implement their own logistical plans for their production capacity and consumer intensity.

## REFERENCES

[1] Henrique Luiz Corrêa and Lucia Helena Xavier, Concepts, design and implementation of Reverse Logistics Systems for sustainable supply chains in Brazil, Journal of Operations and Supply Chain Management , 6(1) 2013

[2] Patricia Oom do Valle, Joao Menezes, Elizabeth Reis and Efigenio Rebelo, Reverse logistics for recycling: The customer service Determinants, International Journal of Business Science and Applied Management, 4(1), 2009

[3] Marisa P. de Brito, Simme D. P. Flapper and Rommert Dekker, Reverse Logistics: a review of case studies, Econometric Institute Report EI 2002-21, Erasmus University Rotterdam, 2002

[4] Leo Kroon and Gaby Vrijens, Returnable containers: an example of reverse logistics, International Journal of Physical Distribution and Logistics Management, 25(2), 1995, 56-68

[5] Aarón D. Castillejo T and Fredrik Stensson, Returnable Plastic Packaging Flow in The Automotive Industry, Master's Thesis E 2011:061, Department of Technology Management and Economics, Chalmers University of Technology, Sweden, 2011

[6] Reza A. Maleki and Jonathan Reimche, Managing Returnable Containers Logistics - A Case Study, International Journal of Engineering Business Management, 3(2), 2001, 1-8

[7] Mikko Kärkkäinen, Timo Ala-Risku and Marianna Herold, Managing the rotation of reusable transport packaging – a multiple case study, 13th International Working Seminar on Production Economics, International Working Seminar on Production Economics, 2004

[8] Isabel Fernández Quesada, The concept of reverse logistics. a review of literature, Annual Conference for Nordic Researchers in Logistics, NOFOMA-03, Finlandia, 2003

[9] Dale S. Rogers and Ronald S. Tibben-Lembke, Going Backwards: Reverse Logistics Trends and Practices, Reverse Logistics Executive Council, 1998

[10] Magdalena Graczyk and Krzysztof Witkowski, Reverse logistics processes in plastics supply chains, Total logistic management, 4, 2011, 43-55

[11] Brandon Kuczenski and Roland Geyer, PET bottle reverse logistics—environmental performance of California's CRV program, International Journal of Life Cycle Assess, 2012

[12] Chee Wong, A study of plastic recycling supply chain, A research project on University of Hull Business school and logistics institute, UK, 2010

[13] Yukie Umeda, Hiroshi Tsukaguchi and Yan Li, Reverse logistics system for recycling: Efficient collection of electrical appliances, Proceedings of the Eastern Asia Society for Transportation Studies, 4, 2003

[14] Georgiadis.P and Vlachos.D, Decision making in reverse logistics using system dynamics, Yugoslav Journal of Operations Research 14(2), 2004, 259-272

[15] Dennis W. Krumwiede and Chwen Sheu, A model for reverse logistics entry by third-party providers, Omega- International Journal of Management Science, 30, 2002, 325 – 333

# AN ACCESS CONTROL MECHANISM WITH PRIVACY PROTECTION FOR RELATIONAL DATA

**Anusha Hiremath[1], Shrikanth Athanikar[2]**

*[1]PG Scholar, [2]Assistant Professor, Dept of CSE, KLE DR MS Sheshagiri College of Engineering and Technology, Belgaum (India)*

## ABSTRACT

*Privacy Preserving Data mining (PPDM) is the new territory of examination that studies the reactions of learning mining systems on people and associations security. With the advancement of data mining innovation, an expanding number of information can be mined out to uncover some potential data about client. While this will prompt an extreme issue, which is clients' protection may be damaged effectively. The objective of privacy preserving is to mine the potential significant information without spillage of sensitive records. An Access Control Mechanism with privacy protection has been proposed in this paper. The access control mechanism characterizes the authorization to get to the asked for traits while the privacy protection mechanism is to conceal the sensitive records. The idea of imprecision bound has been utilized to fulfill the privacy prerequisites.*

***Keywords: Sensitive Attributes, Privacy, Access Control***

## I. INTRODUCTİON

Associations gather and dissect purchaser information to enhance their administrations. Access Control Mechanisms (ACM) are utilized to guarantee that just approved data is accessible to clients. Nonetheless, sensitive data can even now be abused by approved clients to bargain the security of buyers. The idea of privacy protection for sensitive information can require the authorization of security arrangements or the assurance against personality divulgence by fulfilling some security prerequisites. Privacy-preservation from the anonymity aspect has been investigated. The sensitive data, even after the removal of identifying attributes, is still helpless to connecting assaults by the approved clients. This issue has been concentrated on broadly in the territory of small scale information distributed and security definitions, e.g., k-anonymity, l-diversity, and variance diversity. Anonymization calculations use concealment and speculation of records to fulfill security necessities with negligible contortion of smaller scale information. The The anonymity techniques can be utilized with an access control mechanism to guarantee both security and protection of the sensitive data. The security is attained to at the expense of exactness and imprecision is presented in the approved data under an access control policy.

The idea of imprecision bound is utilized for every authorization to characterize a limit on the measure of imprecision that can be endured. Existing workload aware anonymization systems minimize the imprecision total for all inquiries and the imprecision added to every consent/question in the anonymized miniaturized scale information is not known. Making the security prerequisite more stringent (e.g., expanding the estimation of k or l) brings about extra imprecision for inquiries. The issue of fulfilling precision imperatives for individual

authorizations in an arrangement/workload has not been mulled over anytime recently. The heuristics proposed for precision compelled protection saving access control are additionally important in the setting of workload-mindful anonymization. The anonymization for persistent information distributed has been mulled over in writing. The emphasis is on a static social table that is anonymized just once. To epitomize the methodology, part based access control is accepted.

The paper is organized as follows: section 2 contains the related work, section 3 contains the background and section 4 contains the proposed work.

## II. RELATED WORK

The first and most closely related work is the oracle's virtual private database model (VPD) [1]. VPD is a collection of fine grained access control enforced by the server along with the secure application context in the oracle9i database server. It enables to create security policies to control database access at row and column level. It also provides the users a flexible mechanism to build the applications that enforce the security policies they want enforced i.e. only where such control is necessary. VPD enforces the security at a fine level of granularity directly on database tables, views or synonyms. There is no way to bypass the security since the policies are attached directly to the database objects and these policies are automatically applied when the user access the data. Whenever the user accesses a table, a view or a synonym which is protected with an oracle virtual private database policy, oracle database dynamically modifies the SQL statement of the user. This modification contains a WHERE condition (called as a predicate) returned by the function implementing the security policy. Oracle Database modifies the statement dynamically, transparently to the user, using any condition that can be expressed in or returned by a function.VPD offers benefits such as lower cost of ownership, elimination of application security problem, application transparency, and new business opportunities. On the other hand it has several disadvantages: limited scope control, limited predicate size, and repetitive execution.

The limitations of VPD are addressed in [2] which specifies the access control using authorization views. An authorization view is a traditional relational view or parameterized view. A parameterized authorization view is an SQL view which makes use of parameters such as user-id, user-location etc. they proposed two models: Truman model and non-Truman model. Truman model provides each user with a personal and restricted view of complete database. Queries submitted by users will be modified transparently so that the user does not see anything more than his view and the returned answer is correct with respect to the view. A parameterized authorization view is defined for each relation of the database by the DBA. This view defines that all users can have access from this database relation. Transparent modification of user query is carried out by substituting each relation by corresponding parameterized view.

Limitations of virtual private database and Truman model motivated the authors to develop non Truman model. In this model a query is undergone into a validity test, which if it fails the query is rejected and a notification is sent to the user. Otherwise the query is allowed to execute without modification. Under this model a user query is said to be valid if it can be answered only using the information contained in the authorization view. User can write queries against the database relations. The DBA can create several authorization views and any of those views can testify for the validity of the user's query. The non-Truman model guarantees correctness, but it requires powerful query inferencing mechanism. In general such inferencing mechanisms are not decidable and

the query accepted by one database can be rejected by another. Such a kind of unpredictability is undesirable for applications.

Further, the predicate based fine grained access control has been proposed in [3] by S. Chaudhuri et.al. The authors proposed a model for fine grained authorization which is based on adding predicates to authorization grants. This model supports predicated authorization to specific columns, cell level authorization with nullification, authorization for function/procedure execution and grants with grant option. The authorization model proposed by the authors extends the authorization models of the SQL: 2003 standard and add the several new components such as a user context, authorization predicates and query defined user groups.

Access control with privacy mechanisms have been studied by S. chaudhuri in [4]. They proposed a hybrid architecture which combines (a) a set of authorization predicates which restricts each user to only a subset of data (b) a set of noisy views that exhibits the deranged collective information over data which are not accessible through the authorization predicates. Noisy views are the abstraction to integrate privacy mechanisms in traditional database. They use the definition of differential privacy proposed in [5]. Privacy requirements in terms of k-anonymity has been defined in [6]. In [6] they analyzed the k-anonymity and explored its limitations latter they proposed a safe k-anonymity algorithm.

Workload aware anonymization was studied in [7]. They used a simple language for describing a family of target workloads and a suite of algorithms for incorporating these workloads into anonymization process. The authors discussed three techniques: classification and regression, selection, and aggregation and summary statistics.

Existing literature on workload aware anonymization focuses to minimize the overall imprecision for a given set of queries. However anonymization with imprecision constraints for individual queries has not been studied been studied before. The imprecision definition of [12] has been followed and introduced the constraint of imprecision bound for each query in a given query workload.

## III. BACKGROUND

Privacy definitions are briefly explained in this section. For a given relation R= {a1, a2… an} where ai is an attribute, R* is the anonymized table of the relation R. Relation R is assumed to be a static table which includes the following types of attributes.

Identifier Attributes: These are the attributes which individually identifies the user. E.g. name, social security number etc.

Quasi Identifier (QI) Attributes: Attributes, e.g., gender, zip code, birth date, that can potentially identify an individual based on other information available to an adversary.

Sensitive Attributes**:** Attributes, e.g., disease or salary, that if associated to a unique individual will cause a privacy breach.

## IV. PROPOSED WORK

This section is divided to three sub sections. Architecture Design, Access control mechanism: which includes brief explanation about various access control mechanisms, the mechanism which has been adopted in our work and a key generation algorithm has been proposed. The last section is privacy protection: this includes brief about how data will be protected from leakage and also a matrix encryption algorithm has been proposed.

### 4.1 Architecture Design

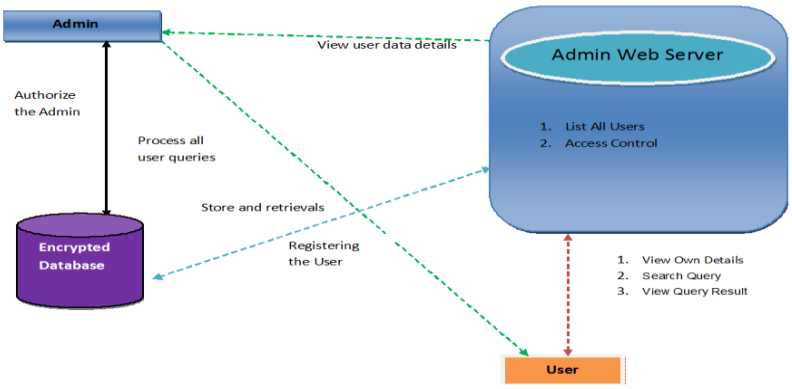



**Fig 1. Architecture Design**

Above fig 1 shows proposed architecture where each user registers with the system then searches for query by providing bound (attribute based on which query needs to be searched). The administrator receives the query and checks for the bound. If the bound is not an identifier attribute then he provides access control for that attribute. Otherwise the query will be denied and the user is intimated about the denial. Privacy is maintained by encrypting the entire database.

### 4.2 Access Control Mechanism

Access control is the process of mediating every request to resources and data maintained by a system and determining whether the request should be granted or denied. An imperative necessity of any data administration framework is to secure information and assets against unapproved revelation (secrecy) and unapproved or ill-advised alterations (integrity), while in the meantime guaranteeing their accessibility to authentic clients (no denials-of-service). Authorizing security accordingly obliges that each access to a framework and its assets be controlled and that all and just approved gets to can occur.

The various types of access control mechanisms [8] are discussed below:

### 4.3 Discretionary Access Control

This is the traditional access control in which user has the complete control over all the programs. DAC is based on giving access to the user on the basis of user identity and authorization which is defined for open policies. DAC owns and executes and also it determines permissions to the particular user to the object. DAC policies considers the access of users to the object which is based on the user's identity and authorization that specifies for each user's access method and object that is requested by user. Each individual request to access an object that has been checked. In DAC access method flexibility will be good. In this method most of the authorization is specified explicitly and also authorizations of individual user is closed. And also when authorizations are open then it is said to be open policies. DAC consists of access rules and access attributes .The access attributes allows the system to define several distinct level of authorization, and the access rules provide the mechanism for the cloud to prevent unauthorized access of sensitive information. DAC provides controlled sharing of objects among various subjects. DAC is said to be the mechanism of "who can access what". In DAC the owner of an object can choose to grant access permissions to other users. Access control list is associated with each object's file system. A simple form of Discretionary access control can be file passwords and giving access to

the authorized users. DAC mainly deals with the following that are Inheritance of permissions, User-Based Authorizations, Auditing of System Events, and Administrative Privileges.

### 4.4 Mandatory Access Control

Mandatory access control is based on the access of objects to number of subjects. Mandatory access control is mainly based on the security level. In this individual cannot change the access. Traditional MAC mechanism is mainly coupled with some security consideration. This follows the following two principles. Those are, read down (users current security level must dominate the access of the object being read) and write up (users current security level must dominate the access of the object being write) MAC based on the classification of objects and subjects present in the cloud environment. Access to a particular object is allowed only if some relationship is satisfied. Each object and subject present in cloud environment assigned some security level. This security level helps to identify the current access state of the object. Security level associated with user also called clearance. MAC used to protect network and file system, block users from accessing without appropriate authorization. In MAC the users will not be permitted to change the access control and its security level. MAC label is said to be security attribute which may be applied to subjects and objects throughout the system.

### 4.5 Role-Based Access Control

In role based access control access decisions are based on the individual's roles and responsibilities within the cloud environment. It formulates the user's access to the system based on the activities that the user has been executed in the cloud. It requires the identification of roles of users on the system. Role can be set of objects or actions associated with the subject. Role may vary depends on the user's priority. RBAC provides the web based application security. Roles are assigned based on the particular cloud organizational structure with their security policies. Each role in the organization's profile includes all authorized users, commands, transaction and allowable information access. Roles can be assigned based on the least privilege. These identified roles can be transferred and used based on the appropriate procedures and security policies. Roles can be managed centrally. RBAC implemented in three ways based on the design constraints that are, RABC0, RBAC1, RBAC2, RBAC3. RBAC0 is based on the least privileges and separation of roles. It does not contain hierarchy and permissions to the particular object is assigned directly. RBAC1 is based on the use of hierarchies and RBAC2 is based on the hierarchy within the RBAC1. RBAC3 is based on the both constraints and hierarchy. RBAC allows users to execute multiple roles at the same time and roles are the useful approach to organizations such as cloud, grid and peer to peer environment. In some cases the only one role can be assigned to one user and it recognize the same roles to other users jointly. After the DAC and MAC Mechanism RBAC has been proven as the efficient access control mechanisms. So securing information on the cloud is similar to securing data on the web. RBAC on the web is user pull architecture. RBAC can be used to provide service and assigns roles to each user's based on the user identity and its role based on the execution environment in cloud. RBAC on the web is implemented with server pull architecture. RBAC permissions are associated with roles and users are assigned to appropriate roles. System administrators only can be able to create roles and granting permissions to those roles. Without RBAC it is difficult to determine what permission has been assigned to which user.

### 4.7 ABAC (Attribute Based Access Control)

ABAC is attribute based access control normally considers identification, authentication, authorization and accountability. In attribute based access control the attributes are considered based on the user's request and the type of access user wish to access and the needed resources of user. ABAC is more secure and flexible and scalable and it provides hierarchical structure. Set of user attributes will be maintained individually.

#### 4.7.1Methodology Used

In our work ABAC Access control mechanism has been used wherein user requests admin to access the attribute of his choice. Administrator will be having the right to decline the request. He will do so if the requested attribute is identifier attribute. Here access is provided by a key through which user can access the attribute. One time key will be provided if the user wishes to access a particular attribute more than once. Key generation algorithm [9] has been used which is as below:

```
Function token generation
Output: key
Begin
  Arr= generate ();
  Print (arr);
 End function
Function generate
Begin
  For i= 0 to 4 step by 1
    C= random character;
     D= random number;
     Key= concatenate (C, D);
Return st;
End function
```

Once the user is provided with the key he can search for a query by providing the bounds and the disease which he want to search. Bounds are the limit for which the result is given. In base paper administrator will be providing bounds which is not convenient for the user. In our work user is allowed to provide the bounds. After getting the bounds and the disease suffix array algorithm has been used for searching [10] which is given below:

```
Function suffixArray
Input: string to be found
Begin
 Initialize suffixarray;
 For i=0 to s.length () step by 1
 Find index;
 Find substring using index
 Assert s.substring (index).equals (suffix.select (i));
 Find rank;
 If (i==0)
   Print (I, index, rank, Substring);
```

```
Else
  Find lcp;
  Print (I, index, lcp, rank substring);
End function
Function rank
Input: string
Begin
  Lo= 0, hi= suffixarray.length-1;
  While (lo < = hi)
     Mid= lo+ (hi-lo)/2;
     Cmp= compare (query, suffixarray [mid]);
     If (Cmp<0) hi= mid-1;
     Else if (Cmp >0) lo= mid+1;
     Else return mid;
End function
Function compare
Input: query, suffix
N= Math.min (query, length);
For i= 0 to N step by 1;
   If (query. (CharAt (i) < suffix.charAt (i)) return -1;
   If (query. (CharAt (i) > suffix.charAt (i)) return +1;
Return query.length () – suffix.length ();
End function
Function lcp
Input: i
Begin
  If (i< 1|| i>= suffixarray.length ())
   Return lcp (suffixarray [i], suffixarray [i-1]);
End function
Function lcp
Input: suffixs, suffixt
N= Math.min (s.length (), t.length ());
For i=0 to N step by 1
  If (s.charAt (i)! = t.charAt (i))
 Return i;
Return N;
End function;
```

Suffix arrays are essentially a representation of the lexicographic order of all suffixes of a string. They can be constructed very efficiently in terms of time and space. The use of suffix arrays makes string mining scalable to large databases.

Input to the algorithm is the string to be matched. It first initializes the suffixarray with the string and finds the substring. Then finds the index computes rank and lcp and then prints the rank, lcp and substrings.

### 4.8 Privacy Protection

Information becomes sensitive when they are specific to a small number of individuals. Data mining, on the other hand, typically makes use of information shared by some minimum number of individuals to ensure a required statistical significance of patterns. As such, sensitive information are to be protected from any leakage. In our work the sensitive information will be encrypted and stored in the database so that even if the database is hacked no one can decrypt the data.

Matrix encryption and decryption algorithm [11] is used for encryption.

#### 4.8.1 Matrix Encryption and Decryption Algorithm

```
Function encrypt
Input: msg, key
Output: cipher text
Begin
   len= key.length ();
   x= 80/len;
  Ckt= 0;
  For j=0 to x step by1
      For i=0 to len step by 1
        Matrix[i] [j] = (char) (46+ckt);
        Ckt++;
  For i= 0 to msg.length () step by 1
     For k= 0 to x step by 1
        For j= 0 to len step by 1
           If (msg.charAt (i) == matrix[j] [k])
              Out+= key.charAt (i) +k+j;
   Return out;
End function
```

Input to the algorithm is the message to be encrypted and globally accepted key. The algorithm first constructs the matrix of numbers, alphabets and symbols (starting from ASCII character 46). It then matches each character of message against the characters in the matrix and the corresponding row and column numbers are concatenated with the corresponding character of the key. This forms the cipher text.

Below is the algorithm used for decryption. Input to the algorithm is the cipher text to be decrypted and the same key which was used for encryption. First it will construct the matrix in the same way as in encryption algorithm. It then extracts the row and column number from the cipher text which is used to search the corresponding character from the matrix. The result will be concatenated with previously computed result and finally the original message will be constructed.

```
Function decrypt
Input: cipher text, key
Output: msg
```

```
Begin
   Len= key.length ();
   x= 80/len;
  Ckt= 0;
  For j=0 to x step by1
      For i=0 to len step by 1
       Matrix[i] [j] = (char) (46+ckt);
       Ckt++;
   a=1;
   b=3;
   loop_len= cipher.length/3;
   While (count<loop_len)
      Asd [count] = cipher.substring (a, b);
      t1= Integer.parseInt (asd [count].charAt (0));
      t0= Integer.parseInt (asd [count].charAt (1));
      Out+=matrix [t0] [t1];
      a= a+3;
      b= b+3;
      Count++;
    Return out;
End function
```

## V. CONCLUSION

An access control mechanism with privacy protection framework has been proposed. Access control mechanism allows only authorized queries. In privacy protection the data is encrypted and stored in the database. In this framework user is allowed to provide the bounds for searching the query which is efficient.

## REFERENCES

[1] K. Browder and M. Davidson, "The Virtual Private Database in oracle9ir2," Oracle Technical White Paper, vol. 500, 2002.

[2] S. Rizvi, A. Mendelzon, S. Sudarshan, and P. Roy, "Extending Query Rewriting Techniques for Fine-Grained Access Control,"Proc. ACM SIGMOD Int'l Conf. Management of Data, pp. 551-562, 2004.

[3] S. Chaudhuri, T. Dutta, and S. Sudarshan, "Fine Grained Authorization through Predicated Grants," Proc. IEEE 23rd Int'l Conf. Data Eng., pp. 1174-1183, 2007.

[4] S. Chaudhuri, R. Kaushik, and R. Ramamurthy, "Database Access Control & Privacy: Is There a Common Ground?" Proc. Fifth Biennial Conf. Innovative Data Systems Research (CIDR), pp. 96-103, 2011.

[5] S. Chaudhuri, R. Kaushik, and R. Ramamurthy, "Database Access Control & Privacy: Is There a Common Ground?" Proc. Fifth Biennial Conf. Innovative Data Systems Research (CIDR), pp. 96-103, 2011.

[6] N. Li, W. Qardaji, and D. Su, "Provably Private Data Anonymization: Or, k-Anonymity Meets Differential Privacy," Arxiv preprint arXiv:1101.2604, 2011.

[7] K. LeFevre, D. DeWitt, and R. Ramakrishnan, "Workload-Aware Anonymization Techniques for Large-Scale Datasets," ACM Trans. Database Systems, vol. 33, no. 3, pp. 1-47, 2008.

[8] Punithasurya K, Jeba Priya S, "Analysis of Different Access Control Mechanism in Cloud", International Journal of Applied Information Systems, Vol 4- No 2, September 2012.

[9] N Ajlouni, A El-Sheikh and A Abdali, "A New Approach in Key Generation and Expansion in Rijndael Algorithm", The International Arab Journal of Information Technology, Vol 3, No 1, January 2006.

[10] M. I. Abouelhoda, E. Ohlebusch, and S. Kurtz. "Optimal exact string matching based on suffix arrays", In Proc. 9th Symposium on String Processing and Information Retrieval, volume 2476 of LNCS, pages 31-43. Springer, 2002.

[11] B Acharya, G Rath and Sarat K P, "Novel Modified Hill Cipher Algorithm", Proceedings of ICETAETS 2008 pages 126-130.

[12] Zahid P, Walid G. Aref, Senior Member, IEEE, Arif G, Fellow, IEEE, and Nagabhushana Prabhu, "Accuracy-Constrained Privacy Preserving Access Control Mechanism for Relational Data", IEEE transactions on Knowledge and data Engineering, VOL. 26, No. 4, April 2014.

# COMPLEX KEYWORD QUERİES SEARCH AND PREDİCTİON OVER DATABASES

**Nikita D Pattanshetti[1], Prof. Savita Bakare[2]**

*[1]PG Scholar, [2]Assistant Professor, Dept of CSE,*
*KLE Dr. M S Sheshgiri College of Engineering and Technology (KLECET),*
*Belagavi, Karnataka, (India)*

## ABSTRACT

*Users query in search engines using keywords and it's a widely used form of querying. In response to a user need, or query, the task of an Information Retrieval system is to retrieve useful information from a large repository of data. Search engines usually do keyword matching only. As long as the query word is present in the document, it is fetched and presented to the user as the result of the query. If a user query is general and ambiguous, the query engine finds the query as a complex query as it cannot extract the exact result which the user expects. Search engines do not identify the needs behind the query hence most of the relevant documents are not retrieved. The difficulty of this task will be affected by various factors, relating to the system or algorithms used, to the properties of the data to be retrieved, or to the inherent difficulty of the user's information need. The effect of these factors upon retrieval performance is often referred to as query difficulty or complexity and the query is said to be a complex query. The aim of this paper is to analyze and study about the efficient query retrieval and prediction over the databases. In this approach, the Structured Robustness (SR) score algorithm is implemented to predict the effectiveness of the keyword queries over databases. Query reformulation method is used where the search engine helps the user to reformulate their query by adding more specific keywords to it. The expected outcome effectively predicts the hard queries on structured data over databases. The ranking quality of the results provides a good user satisfaction..*

***Keywords: Keyword Query, Keyword Query Interfaces, Query Performance, Structured Robustness (SR) Score.***

## I. INTRODUCTİON

Data mining is the process of extracting useful information from large volumes of data. It is also known as Knowledge Discovery in Databases (KDD) and it is used for data extraction from various databases such as relational databases, object-oriented databases, data warehouses, transactional databases etc. With the help of data mining tools we can predict behavior and future trends and make knowledge-driven decisions.

With the growth of the internet and enormous data, it is difficult for the user to get relevant documents. Information Retrieval (IR) system retrieves information from a large repository of data in response to a user query. Given a set of documents in the database and a query given by user, subset of documents relevant to the query is to be retrieved in any Information Retrieval system. In Relational databases, the data is commonly searched using Structured Query Language (SQL). Information needed to answer a keyword query is often split across the tables/tuples. If the user knows the schema of the database he can form suitable query for his needs.

In case of online databases, user usually does not have detailed knowledge of schema or query languages. Hence the desired results are not obtained.

Keyword Query Interfaces (KQI) is the type of computer-human interface used for selecting the required data. Keyword queries on databases are used to provide easy access for the data to be searched. But, these data have low ranking quality in real world scenario [1]. In response to a user's query the Search engines usually do keyword matching and return ranked list of all the documents containing the keywords specified in the query. The relevant documents may not be retrieved and/or retrieved instances may not be relevant (i.e. low precision and/or recall).

Keyword search provides an alternative and easy way of querying in relational databases. One important advantage of keyword search is users need not have prior knowledge about the structures of the underlying data or the knowledge of complex structured query languages (e.g., SQL) for querying their information needs. In the event that the query is general, it is hard to recognize the specific record on which the user is interested. The users are required to filter through a not insignificant rundown of off-subject reports. The queries which are hard to answer effectively by the Information Retrieval (IR) system are called hard queries or complex queries. Thus, the query engine should identify and search the desired attributes associated with each term given in the query. Thus, it is important to distinguish such queries that are liable to have low positioning quality with a specific end goal to enhance the user fulfillment level.

The topic of how to find data of interest of user in the World Wide Web is raised by the Web Search Problem [2]. Majority of queries requested to the Internet search engines by the users are general, few words in length. Low quality searches are many and therefore methods of dealing with the results of such queries are needed. One method is filtering of the ranked list of documents, varying from simple pruning techniques to advanced Artificial Intelligence algorithms. Although they limit the total length of the ranked list, it is difficult for users to locate the specific documents they searched for. The search engines can also help the users to refine their query by adding more specific keywords to it.

Query engine must assign each query words to schema elements in the database. In this regard, this paper we study the properties of complex queries over databases and proposes a method to detect such queries. The structure of the data in the given database is used to study about the degree of the complexity of the query. The method used predicts the degree of the complexity of a query efficiently. Structured Robustness (SR) score [1], measures the difficulty of a query based on the differences between the rankings of the same query over the original and noisy (corrupted) versions of the same database. Finally, thresholding approach is used to define query difficulty metric.

The rest of this paper is organized as follows. Section 2 describes related work. In section 3, we study the different prediction methods. Section 4, defines the properties of complex queries. In section 5, we describe the data models used. In section 6, we explain about the main methodology of our work. In section 7, we briefly explain about the result obtained. In Section 8, we summarize the main conclusions of this paper.

## II. RELATED WORK

Predicting the query performance is important for an Information retrieval system. It is studied under different names like query difficulty, query complexity, query ambiguity and sometimes hard query. Existing work on

query complexity estimation can be categorized along three axes [3]. How is query complexity defined, How is query complexity predicted and How is the quality of the prediction evaluated.

### 2.1 Defining Query Complexity

We can define query hardness or complexity in many ways; for example, queries can be inherently complex or can be ambiguous, difficult for a particular collection of data [3]. A query can be complex in a given collection of data if it has more outcomes for single query. For example, Carmel et al. [4] considered collection query hardness by comparing the query difficulty predicted by their method to the median average precision taken over all runs submitted in the Terabyte tracks at TREC (Text REtrieval Conference) for a given query.

### 2.2 Predicting Query Complexity

Cronen-Townsend, Y. Zhou, and B. Croft in [5] introduced the clarity score which effectively measures the ambiguity of the query with respect to a collection. Clarity scores are computed by assessing the information theoretic distance between a language model associated with the query and a language model associated with the collection. They showed that clarity score positively correlates with average precision.

### 2.3 Evaluating the Quality of Query Complexity Predictions

In order to evaluate the quality of a query hardness prediction methodology, the collection hardness of a set of queries is measured and they are compared to predicted values of query hardness. These actual and predicted values are real-valued, and they are typically compared using various parametric and nonparametric statistics.

S. C. Townsend, Y. Zhou, and B. Croft in [5] use clarity score method to predict the query outcome by computing a measure of disorder between a query model and its collection model. The subsequent clarity score measures the coherence of models that are likely to generate the query. A threshold for clarity score is set between predicted queries and acceptable queries and it is validated using TREC (Text REtrieval Conference) data. They showed that clarity score measures the ambiguity of a query and it is positively correlated with average precision in a variety of TREC test sets.

The clarity score method generally assumes that the lengthier is the query, the easier it will be to evaluate. To extend idea of clarity score for queries over databases, domain knowledge about the data sets is required. Empirical studies show that these methods have limited prediction accuracies [5].

Y. Zhou and B. Croft in [6] introduced the notion of ranking robustness. It refers to a property of a ranked list of documents. It indicates how strong the ranking is in the presence of noise in the ranked documents. Robustness score, a statistical measure is used to quantify this notion. Given a query and a ranking function the ranking robustness is calculated by comparing the ranking from the corrupted dataset and ranking from original dataset. They showed that robustness score correlates with query outcomes in a variety of TREC data collections. Query difficulty prediction model for our work falls under this category.

B. He and I. Ounis, in [7] Use a set of predictors which are computed before the retrieval process takes place. IR system scans the files for the query terms and assigns a relevance score to each retrieved document. Some of the predictors are query length, which is number of words in the query, query clarity which defines the unambiguous property of query. The query scope was originally studied in [8] which defines how general or specific is a query. Some of the predictors have significant correlation with query performance. Therefore, these predictors can be applied in practical applications.

Since the predictors are generated and studied before the retrieval of results takes place, therefore domain knowledge about the data sets used is required. It requires finding a similarity function between entities that are about a similar topic. The domain knowledge and understanding users' preferences is required to understand the similarity function.

Some use IDF-related (inverse document frequency) features as predictors. He and Ounis in [9] proposed a predictor based on the IDF of the query terms. IDF-based indicators demonstrated some moderate relationship with query outcome. They estimated the clarity score model by the term frequency in the query. They also used the notion of the query scope, which is quantified as the percentage of entities that contain at least one query keyword in the collection of documents. These predictors usually do not take the retrieval algorithms into account and thus are unlikely to predict query performance well. These kinds of predictors rely on the amount and characteristics of available training.

G. Bhalotia, A. Hulgeri, C. Nakhe, S. Chakrabarti, and S. Sudarshan, in [10] used a system which enables keyword search on relational databases considering both data and schema browsing called BANKS. It is an acronym for Browsing ANd Keyword Searching. It uses backward expanding search algorithm for finding and ranking query results. Using BANKS users can extract data without the need for writing complex queries and without any knowledge of the schema. Database is viewed as a graph where the nodes are the database tuples and the edges are application specific relationships. A user gets his information by typing a few query keywords, following hyperlinks provided, and interacts with the displayed results.

A drawback of these approaches is that a graph of the database tuples must be materialized and maintained. Once the data graph has been built, the structural information provided by the database schema is ignored. This method can be relatively slow, since a large number of tuples may be defined to be relevant to the keyword.

## III. PREDICTION METHODS

Performance Prediction of complex queries is important for many reasons. From the user perspective, it gives feedback to the user to direct his search. From the perspective of retrieval system, it is important to improve the consistency and use alternative retrieval strategies to improve the performance of retrieval systems. The methods to predict complex queries can be broadly categorize into two groups: pre-retrieval and post-retrieval methods [1].

1. Pre-retrieval methods: These methods are used to predict the difficulty of a query without considering its outcomes. The properties of the terms in the query are used to measure the specificity or ambiguity of the query to predict its difficulty [11]. These methods assume that the more are the query terms, the easier the query will be. Actual studies indicate that these methods have limited prediction accuracies [5].
2. Post-retrieval methods: After query computation these methods uses the outcomes of a query to predict its difficulty and generally falls into one of the following categories.

- Clarity-score based methods consider few topics of the document which are frequently searched by user [7]. It deals with most searched topic of the document and retrieves the result from it. The degree of difficulty for query is less if result of user query is found on most searched documents otherwise high.
- Ranking score based methods defines the ranking score of the document returned by retrieval system for a user query. It estimates the similarity between a query and the document and defines the difficulty of a

query by the difference between weighted entropy score of top ranked result and the score of other documents.

- Robustness-based methods defines how robust is the query over specific document. This method defines the degree of difficulty of query by considering the robustness of the ranking over two versions of data, original version and corrupted version, and compares the top k results of same query over these two versions [12]. Degree of difference between these results defines the degree of hardness of query.

Some methods use machine learning techniques to study complex queries and its properties and then predict their hardness [13]. These methods are effective, if large amount and quality of data are available which are normally not available for many databases.

## IV. PROPERTIES OF COMPLEX QUERIES ON DATABASES

The keyword queries which are difficult to answer correctly by the Information retrieval systems are called complex keyword queries or hard queries. There are basically three sources of difficulty for answering a query over a database [1]. They are as follows:

### 4.1 A Entity Matching Many Terms

If one or more entities in the database match the terms in a query, the query is said to be less specific and it is harder to answer correctly. For example: Consider user submits query Q1: JOEL in IMDB (Internet Movie Database) database. The keyword query interface must resolve desire "JOEL" that satisfies user's information need. If there is more than one person named JOEL in the IMDB database then it will be hard to predict the correct answer by the retrieval system. As opposed to query Q1, query Q2: KIM matches the smaller number of people in IMDB, so it is easier for query interface to return relevant answer.

### 4.2 Attribute Level Ambiguity

Each attribute explains a feature of an entity. If a query terms matches different attributes in its database then it will have a more set of answers and hence it has higher attribute level ambiguity. For example: Q3: GODFATHER, contains in title and the distributor of IMDB dataset. Keyword query interface must find out the desired attribute for GODFATHER to find correct answer. Answer for the query Q4: FURIOUS does not match with any instance of attribute distributor, so keyword query interface easily predict the relevant answer for this query.

### 4.3 Entity Level Ambiguity

Each entity set contains the data about different entities. Hence, if a query matches entities from two or more entity sets, it will have higher ambiguity at entity level. For example: IMDB has two entity sets mainly one is 'Movie', which contains the information about movie and other is 'Person', which contains the information about people who are involved in making movies. Consider the query Q5: 'MARRIED', the both Movie and Person contains MARRIED word, so the query becomes complex to find relevant answer as keyword query interface does not know whether user is interested in people who are MARRIED or the movies containing MARRIED word. In case of query Q6: COMEDY SUSPENSE matches with the entities of Movie entity set. So it will be less difficult for keyword query interface to find relevant answer as compared to query Q5.

## V. DATA MODELS

Database is modeled as single entity set called 'Movie'. Each entity has a set of attributes that describes the entity. Each attribute contains attribute values. The database fragment is as shown in Fig 1. For instance, title and director are the attributes and 'Heartless' and 'Aditya chopra' are its attribute values respectively. The entities, which are the movie titles, are stored using set of normalized relational tables.

The model below is widely used in works on entity search [14]. It has an advantage that it can be easily mapped to both XML and relational data.

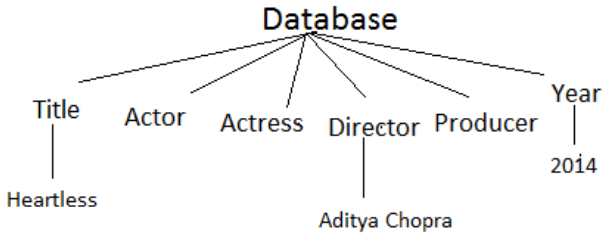



**Fig 1: Database fragment**

Keyword query Q is a set of keywords given by a user. An entity E from the database is the answer to Q if and only if at least one of the attribute values contains those query words. Given a database DB and a query Q, the retrieval function g(E, Q, DB) returns the relevance of entity E in DB to query Q. A keyword search system returns a ranked list of matched entities in DB called L(Q, g, DB) where entities E are placed in decreasing order of the value of g(E, Q, DB).

## VI. METHODOLOGY

E.Mittendorf in [15] has shown that if a retrieval system effectively responses and ranks the results to a user query in a collection of documents, it will likewise perform well for the same query over the corrupted version that contains some errors such as repeated words. In other words, the degree of the complexity of a query is positively correlated with how robust is its ranking over the original and the corrupted versions of the data. This observation is called the Ranking Robustness Principle [1]. It defines how robust is the ranking in presence of noise.

In this paper the database is corrupted by adding repeated attributes. First we define data corruption model for structured data. For that we model a database DB using entities, attributes, and attribute values. The noisy or corrupted version of the database is created by adding repeated attribute values for entities. Given a user query Q and a information retrieval function g, we rank the candidate answers in DB to get the ranked list L. The more results in L the more complex are the query.

To categorize input queries thresholding approach is used. A query is said to be an as "easy" or "complex" query based on certain threshold [5]. This thresholding approach defines a reasonable threshold "t" for a query difficulty metric. If the difficulty metric of the query is above t, it will be considered as a complex query, and the KQI will apply further treatments like the user is asked to provide some description about the query with one or more keywords.

For a easy query with few matched entities, the retrieval system easily fetches the records from the database. For a lengthy and complex keyword query it is hard to get the accurate result over database as the retrieval system cannot identify the desired attributes for the user query. Hence multi-stage processing is required to get the proper answer for the user query.

First, Preprocessing of the data involves removal of stop words and stemming. Stopwords are normally excluded from the set of query terms. It is important to extract only the keywords from the query. Natural candidates for stopwords are articles, prepositions and pronouns. Porter stemmer is a stemming algorithm that is based on the context aware [16]. It is a process for removing the commoner morphological endings from words in English. Its main use is in normalization process which is important when using Information Retrieval systems. For instance "Movies' in a query is preprocessed as 'Movie'. It is used to increase recall without decreasing precision.

After preprocessing the important words of the query are extracted and matched against the database for accurate result. Clustering is a process of forming clusters of similar objects from a given set of inputs. In case of web search results, clustering is a way of organizing the results into a number of easily browse-able groups. For clustering of results obtained from search engine we use basic concepts of algorithm called LINGO [17]. The key idea is to first discover cluster labels and determine the data of the groups based on the labels. Two clusters labels formed are a cluster for an easy query with one or two results and another cluster for complex query for which more than two results are fetched.

Structured Robustness (SR) score is used to measure the difficulty of a query [1]. Structured Robustness Algorithm computes the exact SR score based on the results for a query. Some statistics that can be used are the number of query terms matching all attributes values of the database or number of results fetched for a query.

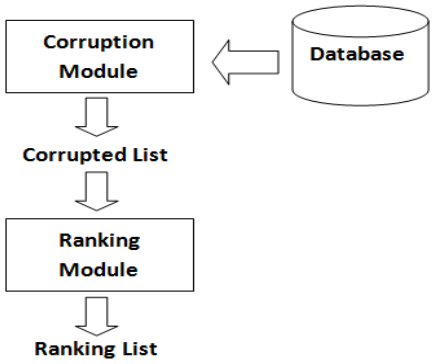



**Fig 2: Structured Robustness Algorithm**

The Structured Robustness algorithm is as shown in Fig 2. SR Algorithm checks for all attribute values in the results for all query terms. The corruption module checks for all the keywords in the database. A corrupted list with all possible keywords matched in database is obtained. With the user description about the query the most appropriate result is fetched and presented to the user. The ranking module ranks the results with most appropriate result at the top of the list.

We use the XML ranking method used in [18], called PRMS- Probabilistic Retrieval Model for Semistructured Data. Given a query Q, PRMS computes the relevance score of entity E based on the weighted linear combination of the relevance scores. It assigns each attribute a specific weight, which specifies its contribution in the ranking score. The algorithm for the same is given below.

Algorithm:

Input: Documents to be ranked.

Output: Ranked list of documents.

For (i=1 to n) Do

Search the documents and compute the results.

Calculate and cache them.

For (j= 1 to n)

Calculate the rank for the given query input.

Display the ranked list of documents.

For a given query, the documents fetched by the retrieval system are ranked according to the user preferences. We calculate rank for the given query input by considering the number of clicks for a document. Document with highest clicks is fetched and presented at the top of the ranking.

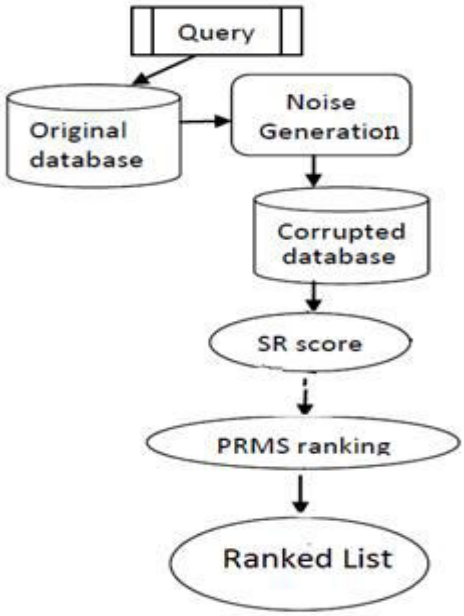



**Fig 3: Work Flow Architecture**

The workflow architecture of the whole idea is as shown in Fig 3. User enters a query according to his information needs. The database is corrupted by generating noise in the database. Noise generation, as mentioned earlier, is done by adding repeated attribute values. The user query is now searched in corrupted database. The keywords in the query are matched against every attribute in the database and appropriate results are fetched. The SR score obtained defines the complexity of the query. Finally using PRMS ranking a ranked list of documents is provided to the user as the search result with most searched entity is listed first. This ranking provides a good user satisfaction as the most appropriate result is present at the top of the list.

## VII. RESULTS

The database used for this work contains 500 records. When a user issues a query, which is not ambiguous, relevant records were fetched by the retrieval system. For a complex query, having difficulty metric above the threshold, the user is asked to provide description about the query with one or more keywords. Based on the keywords provided by the user relevant records were fetched. Numbers of records fetched were less for easy query, where as for complex query which is ambiguous, three or more records fetched with the most relevant

record at the top of the list. The ranking quality of the results provided a good user satisfaction. The algorithm predicts the complexity of queries with fewer errors and in negligible time.

## VIII. CONCLUSION AND FUTURE WORK

This paper focuses on main problem of retrieving appropriate top results for a keyword query and predicting the difficulty level of the query. In this paper, we analyze the properties of complex queries and measure the degree of complexity of a keyword query over a database. SR algorithm is used for difficult keyword queries prediction over databases. We measured the degree of the complexity of a query over a database, using the ranking robustness principle. The framework efficiently predicts the effectiveness of a keyword query.

The future work can be extending this framework to estimate query difficulty by using different entity sets and also on other ranking problems on databases. It can be extended for semi-structured keyword queries with operators and supporting phrases provided by the user.

## REFERENCES

[1] Shiwen Cheng, Arash Termehchy, and Vagelis Hristidis, "Efficient Prediction of Difficult Keyword Queries over Databases" IEEE Transactions on Knowledge and Data Engineering, Vol. 26, No. 6, June 2014.

[2] Eric W. Selberg. "Towards Comprehensive Web Search" Doctoral Dissertation, University of Washington, 1999. IEEE Transactions on Knowledge and Data Engineering, August 2013

[3] J. A. Aslam and V. Pavlu, "Query Hardness Estimation using Jensen-Shannon Divergence among Multiple Scoring Functions," in ECIR, 2007.

[4] Carmel, D., Yom-Tov, E., Darlow, A., Pelleg, D.: What makes a query difficult? In: SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval, New York, NY, USA, ACM Press (2006) 390–397

[5] S. C. Townsend, Y. Zhou, and B. Croft, "Predicting Query Performance," in SIGIR, 2002

[6] Y. Zhou and B. Croft, "Ranking Robustness: A Novel Framework to Predict Query Performance," in CIKM, 2006.

[7] B. He and I. Ounis, "Query performance prediction," Inf. Syst., vol. 31, pp. 585–594, November 2006

[8] V. Plachouras, I. Ounis, G. Amati, and C. J. van Rijsbergen. University of Glasgow at the Web Track: Dynamic application of hyperlink analysis using the query scope. In Proceedings of the Twelth Text Retrieval Conference (TREC 2003), pages 248 - 254, Gaithersburg, MD, 2003.

[9] B.He and I.Ounis. "Inferring query performance using preretrieval predictors", In proceedings of the SPIRE 2004. Pp 43-54, 2004

[10] G. Bhalotia, A. Hulgeri, C. Nakhe, S. Chakrabarti, and S. Sudarshan, "Keyword Searching and Browsing in databases using BANKS," in ICDE, 2002

[11] C. Hauff, L. Azzopardi, and D. Hiemstra, "The Combination and Evaluation of Query Performance Prediction Methods," in Advances in Information Retrieval, 2009

[12] E. Demidova, P. Fankhauser, X. Zhou, and W. Nejdl, "DivQ: Diversification for Keyword Search over Structured Databases," in SIGIR, 2010.

[13] E. Yom-Tov, S. Fine, D. Carmel, and A. Darlow, "Learning to Estimate Query Difficulty: including applications to missing content detection and distributed information retrieval," in SIGIR, 2005

[14] N. Sarkas, S. Paparizos, and P. Tsaparas, "Structured Annotations of Web Queries," in SIGMOD, 2010

[15] E. Mittendorf and P. Schauble, "Measuring the Effects of Data Corruption on Information Retrieval," in SDAIR, 1996

[16] Daniel Waegel "The Porter Stemmer"

[17] Stanislaw Osinski "An Algorithm for Clustering Of Web Search Results"

[18] J. Kim, X. Xue, and B. Croft, "A Probabilistic Retrieval Model for Semistructured Data," in ECIR, 2009.

# ASSEMBLING A WEB-CRAWLER

**Jitali Patel[1], Varun Raval[2]**

*[1]Assistant Professor, [2]Student, Dept of CSE, Institute of Technology, Nirma University, Ahmedabad, (India)*

## ABSTRACT

*Web crawlers are nowadays most important part of the internet. Can we imagine our life without GOOGLE, YAHOO and Bing? The web-crawlers are heart of all those search engines. Besides just web-crawler, many complex technologies such as page ranking system, effective key search techniques, disabling malicious sites and much more are involved in making a search engine. This essay is dedicated for complete understanding of a basic web-crawler and some of its requirements. Keep in mind that for commercial web crawlers, both hardware and software requirements are much more than what is described over here.*

***Keywords: Web Pages, Web Crawler***

## I. INTRODUCTİON

A web crawler is a set of algorithms that scans all the websites available on the internet mainly for the purpose of retrieving and updating information related to a specific topic/s. Web-Crawler is mainly used by search engines for updating and retrieving new information to give accurate results of user search. The first inspiration to make web-crawler in mid 1990s was due to the fact that the number of IP addresses grew more than 1 million and it was a big trouble to get latest information on required topic which was made possible by use of web-crawler, key search and text retrieval techniques. How does a crawler fetch "all" Web pages? Before the advent of the Web, traditional text collections such as bibliographic databases and journal abstracts were provided to the indexing system directly, say, on magnetic tape or disk. In contrast, there is no catalog of all accessible URLs on the Web. The only way to collect URLs is to scan collected pages for hyperlinks to other pages that have not been collected yet. This is the basic principle of crawlers. They start from a given set of URLs, progressively fetch and scan them for new URLs (outlinks), and then fetch these pages in turn, in an endless cycle[1]. New URLs found thus represent potentially pending work for the crawler. The set of pending work expands quickly as the crawl proceeds, and implementers prefer to write this data to disk to relieve main memory as well as guard against data loss in the event of a crawler crash. There is no guarantee that all accessible Web pages will be located in this fashion; indeed[2], the crawler may never halt, as pages will be added continually even as it is running. Apart from outlinks, pages contain text; this is submitted to a text indexing system. In this essay, I will be guiding you on the steps required to make a basic web-crawler.

## II. CRAWLING THE WEB

Initially we start crawling with a set of URLs. Even a single URL can be used to start a crawl. The main steps of any crawler are:

1. Fetching IP address
2. Parsing page to get required information

After fetching IP address, page is downloaded and analyzed. From this URL page, we get URLs that this page is pointing at and this new list of URL will again be converted to list of IP addresses, downloaded and analyzed. This is the general procedure of any web crawler whose individual steps are explained in detail below.

DNS Catch and Prefetch

Any web page URL is composed of two types of information. Host name and page destination within that host. E.g. en.wikipedia.org/wiki/main_page, here en.wikipedia.org is host name and page destination is /wiki/main_page. DNS (Domain Name Server) is used to obtain IP using host name or server name i.e. en.wikipedia.org in our example. Steps to maintain database are described below:

1. All new pages fetched from visited pages are added to a list of unvisited pages
2. As need arises (not enough pages to download), these new pages in sequence will be added to a DNS queue that is maintained locally.
3. This DNS queue will be used by crawler to fetch IP address from DNS server.
4. After fetching IP address, these URLs including IP address and page destination e.g. 208.80.154.224/wiki/cricket will be added to a list from where downloader will download pages.

A desktop PC with 256 MB of RAM and a disk cache of a few GB will be adequate for a caching DNS, but it may help to have a few (say, two to three) of these.

Multiple Concurrent Fetches

To save time and to maximize use of any good processing computer, concurrent fetches can be made using extensive use of multithreading and binding more than one socket to DNS server for retrieval of multiple IP addresses at single time. Speed of IP fetches will be dependent upon the number of sockets allocated, processing speed of computer and speed of internet connection[3].

Parsing page

Parsing page means retrieving needed information from a page like metadata, hyperlinks, text attached to hyperlinks, title of page, etc. Metadata include information like expiry date of that page, last update information, type of contents in that page and similar things. Metadata is used to decide when we should revisit that same page for updating our database. For parsing a page, parsers such as DOM, SAX and STAX can be used. These parsers can download the whole page in html format so that any page can be revisited locally. Parsers are mainly used by web-crawlers to search for hyperlinks in page and add them to fetch list and get as much information from metadata about the page as possible[5].

Link Extraction and Normalization

While parsing a page we encounter two types pf URLs:

1. Absolute URL
2. Relative URL

Absolute URLs are the URL which are in their complete form. E.g. en.wikipedia.org/wiki/cricket_ball. This URL can directly be included in the list of URL for fetching IP addresses. But relative URLs are having just page reference relative to the current host. E.g. /wiki/cricket_ball. These types of URLs are needed to be converted to normal form before adding to fetch list.

Robot Exclusion

Many websites have sensitive information on their pages such as account information and other confidential information and these websites don't want the crawlers to crawl such pages. So, Next step is to check whether the server restricts crawling a URL using the robots.txt mechanism. This file is usually found in the Http root directory of the server (such as en.wikipedia.org/robots.txt). This file specifies a list of pages that crawlers should *not* attempt to fetch. The robots.txt file is meant for crawlers only and does not apply to ordinary browsers. If anyone tries to crawl such pages than our IP address may be restricted by that server.

Eliminate already existed URLs

Each and every time a new URL is encountered while parsing a page, it is checked against the list of URLs whose IP addresses are already fetched. For this reason we are needed to maintain a table which will include a list of URLs whose IP addresses have been fetched. Each and every time IP of any address is fetched it will be added to this table. In this way, we can eliminate already existed URLs. But sometimes crawling already visited pages may prove useful. There two types of web-pages

1. Static
2. Dynamic

This information about any page can be obtained from the "last updated" information in meta tags of that page. If last updated information is later than our last fetch of that page, we must refetch that page to update our database.

After Parsing a Page

After parsing a page it should be removed from the list containing a pages to download and added to a list of already downloaded pages just for sake of maintaining a record of downloaded pages. After parsing a page if number of pages to be downloaded are less than certain threshold value (not enough pages to download) than if process for fetching of IP is waiting for a signal to start fetch, it should be notified[4].

Robots.txt

Many websites have sensitive information on their pages such as account information and other confidential information and these websites don't want the crawlers to crawl such pages. For this reason, all the servers contain a file in their root directory called robots.txt. robots.txt file contain information about pages which are disallowed by the server to crawl. If anyone tries to crawl such pages than our IP address may be restricted by that server.

Fetch at some max rate

It is advisable to be kind to any server in terms of rate of fetching information. Any server imposes a limit on the number of requests from any IP address in a given time interval. This number is dependent on individual servers. Thus, before downloading any page we must check if that page's URL is not in the list of busy servers. If that page is in the list, we must wait until the sockets using that page end their connection and go further leaving that page to download in future. If that page is not in the list, we can stat downloading and should see weather that server has now become a busy server and if so, it must be added to the list of busy servers.

## III. TEXT REPOSITORY

Many web-crawler like to maintain a text repository of crawled web pages especially for the webpages that are static. These local repository will be used for performing offline operations such as page ranking, giving

prompts to user based on user entered keyword, etc. Making a text repository will be beneficial if space constraint is not there, internet bandwidth is more and pages to be downloaded are static.

Conclusion

Following the above mentioned steps in programming language of your choice, you will be able to build a small scale web crawler which is able to fetch DNS addresses at some hundred pages per sec. rate and download as per your internet speed and type of text you want to download. A good multithreaded program involving use of multiple sockets in today's 5th gen computers can help crawl pages at faster rate. In order to go for steps to build a search engine, you will need to add several functionalities like page ranking system, key search algorithms, congestion control techniques and obviously a large data space is required.

## REFERENCES

[1]. Chakrabati, Soumen.Mining The Web. Bombay: Morgan Kaufmann Publishers

[2]. Markov, ZDRAVKO. Data Mining The Web. New Britain: Wiley Interscience Publications

[3]. Liu, Bing. Web Data Mining. Chicago, USA: Springer-Verlag Berlin Heidelberg

[4]. Rungsawang, Arnon, and Niran Angkawattanawit. "Learnable topic-specific web crawler." Journal of Network and Computer Applications 28, no. 2 (2005): 97-114.Zhang, Huaxiang. An online semi-supervised clustering approach to topical web crawlers. Science Direct (2010): Pages 491 - 495.

[5]. Stevanovic, Dusan, Aijun An, and Natalija Vlajic. "Feature evaluation for web crawler detection with data mining techniques." Expert Systems with Applications 39, no. 10 (2012): 8707-8717.

# AUTOMATED IRRIGATION SYSTEM USING WIRELESS SENSOR NETWORKS AND GSM MODULE

**Syed Mubarak[1], Shoukhi Khan[2], Sahana.N[3], Megha B[4], S.Sujatha[5]**

*[1234]U.G.Students, [5A]ssociate Professor, Department of Telecommunication Engineering*

*Cmr Institute of Technology, Bangalore (India)*

## ABSTRACT

*Agriculture is the source of livelihood of majority Indians and it also has a great impact on economy of the country. In dry areas or in places where there is inadequate rainfall or when a farmer is not aware of how to use water when there is too much of water storage or when there is no water storage irrigation becomes tedious. Wireless sensor networks and their application for precision agriculture is an automated irrigation system used to enhance the use of water for agricultural crops. The system consists of a wireless network that is the soil moisture sensor and temperature sensor placed under the soil where plants roots are reached which is a distributed network. The system has a water level sensor which will indicate the presence of water level in tank. A gateway unit manages the information related to sensors which triggers the actuators, and data is transmitted using GSM module. A software application was developed by predetermining the threshold values of soil moisture, temperature and water level that was programmed into an arm controller. The data from GSM is transmitted /received from/to mobile using software application or normal texting mode which optimizes the use of water quantity.*

***Keywords: ARM-7- LPC2148 Controller, Embedded C, Flash Magic, GSM, Keil-Compiler.***

## I. INTRODUCTION

**A**GRICULTURE is the need of most of the Indians livelihood and it is one of the main sources of livelihood. It also has a major impact on economy of the country. A major quantity of water is used for irrigation system and therefore 85% of available fresh water resources are used for yielding agricultural crops. This resource of water will decrease day by day and consumption of water will dominate and increase more than 85% in upcoming half century. This is due to the high growth in population due to this tremendous growth in population there is huge demand for food. Agriculture is the main source for food production. Using science and technology we need to implement a method by which there can be limited consumption of water.

Till date many methods have come into existence where water can be limitedly consumed. A method where monitoring water status and based on status of water whether it is high or low irrigation is scheduled which is based on canopy temperature of plant, which was captured with thermal imaging. Another method is making use of information on volumetric water content of soil, using dielectric moisture sensors to control actuators and save water, instead of the scheduled irrigation at a particular time of day and supplying water only for a specific duration. This above method just opens the valve and supply water to bedding plants when volumetric content

of soil will drop below threshold value. In this paper a use of the second method where sensors are placed and based on that water is supplied to the field and intimated to the farmer using software application.

Wireless sensor networks is also called as wireless sensors and actor network, are distributed spatially autonomous sensors to monitor physical or environmental conditions as temperature, pressure sound, moisture etc. and it co-operatively passes these data via network to the main location. WSN is built of few to several thousand nodes, where each node is connected to sensors each sensor network node has typically several parts: a radio transceiver with an internal/external antenna, a microcontroller, an electronic circuit for interfacing with sensors and an energy source such as battery. It has many applications such as area monitoring, health care monitoring, environmental sensing, earth sensing, air pollution monitoring, forest fire sensing, landslide detection, water quantity monitoring, natural disaster prevention, music technology.

## II. EXISTING SYSTEM

At present there is emerging global water crisis where managing scarcity of water has become a tedious job and there are conflicts between users of water. This is an era where human use and pollution of water resource have crossed the levels which lead to limit food production and low down the ecosystem. The major reason for these limitations is the growth of population which is increasing at a faster rate than the production of food and after a few years this population will sum up to 3-4 billion. Thos growth can be seen in countries which have shortage of water resources and are economically poor. Because of growth in population there is a huge demand to raise food production by 50% in the next half century to maintain the capita, based on an assumption that productivity of existing farm land does not decline. The crop water stress index called as CWSI existed around 30 years ago. This crop water stress index was then integrated using measurements of infrared canopy temperatures, ambient air temperatures, and atmospheric vapor pressure values to determine when to irrigate using drip irrigation. The management of these farms which are in greenhouses will require a data acquisition to be located in each greenhouse and the control room where a control unit is located. These are separated from the production area. At present, the data is transferred using wired communication called field bus. This data is transferred between greenhouses and control room. All the problems related here is presented using CAN and ZigBee protocols.

## III. PROPOSED SYSTEM

Here in this paper an experimental scale within rural areas where there is a huge deployment of irrigation system which is implemented using arm controller and wireless communication. The main of this implementation was to demonstrate that the automatic irrigation system can be used to optimize /reduce water usage. It can also be a photovoltaic irrigation system which consists of a wireless network that is the soil moisture sensor and temperature sensor placed under the soil where plants roots are reached which is a distributed network. The system has a water level sensor which will indicate the presence of water level in tank.

A gateway unit manages the information related to sensors which triggers the actuators, and data is transmitted using GSM module. A software application was developed by predetermining the threshold values of soil moisture, temperature and water level that was programmed into an arm controller. The data from GSM is transmitted /received from/to mobile using software application or normal texting mode which optimizes the use of water quantity.

The communication between sensors and data is through the ZigBee protocol. The receiver unit in this system has a duplex communication which is based on cellular internet interface which is done using GSM/GPRS protocol. This is a packet oriented mobile data service which is using 2G and 3G cellular global system for mobile communication.

## IV. BLOCK DIAGRAM

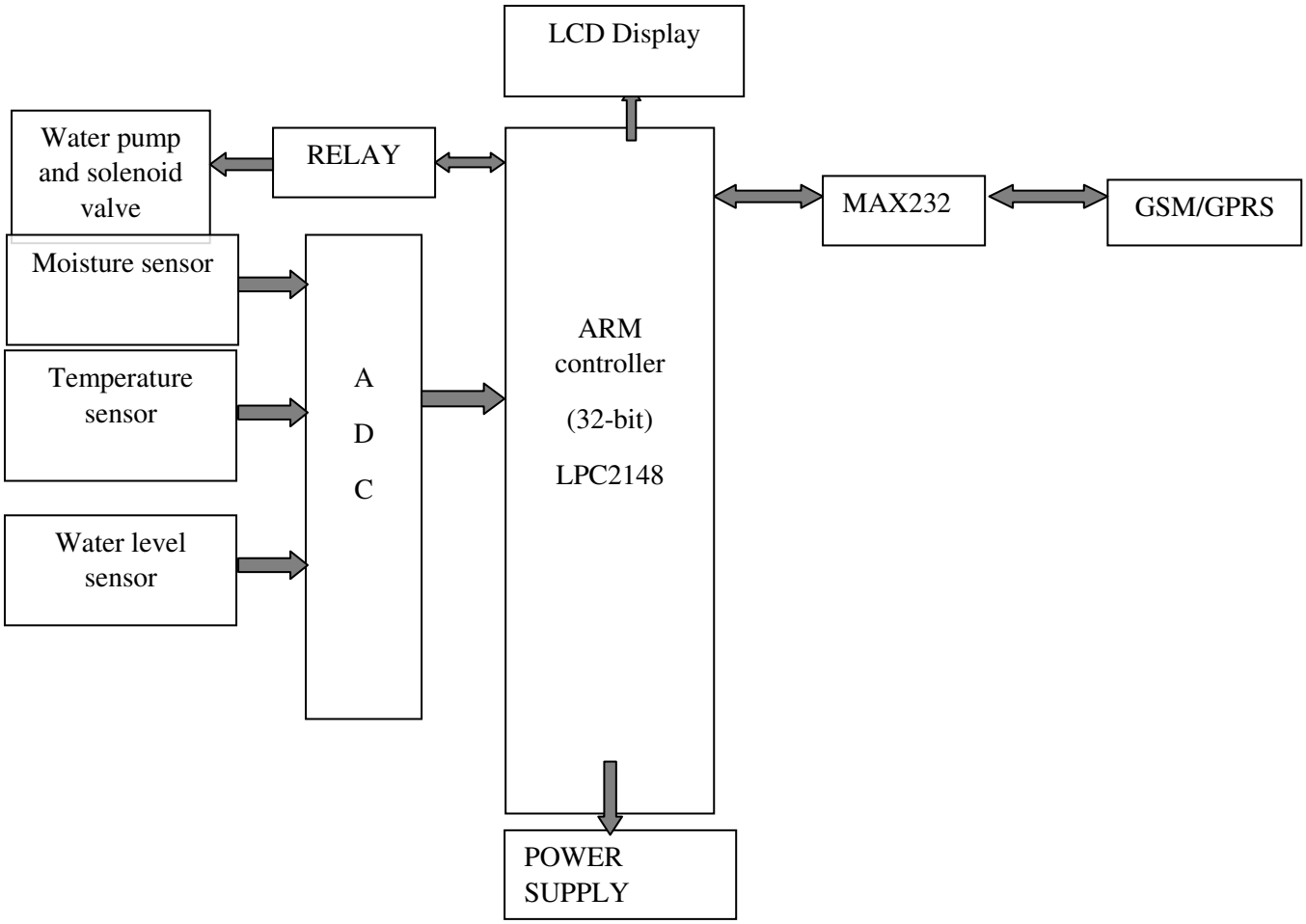



**Fig 2.1 Block Diagram**

ARM is a family of instruction set architectures for computer processors based on a reduced instruction set computing (RISC) architecture developed by British company ARM Holdings. A RISC based computer design approach means ARM processors require significantly fewer transistors than typical CISC x86 processors in most personal computers. This approach reduces costs, heat and power use. These are desirable traits for light, portable, battery-powered devices including smart phones, laptops, tablet and notepad computers, and other embedded systems. A simpler design facilitates more efficient multi-core CPUs and higher core counts at lower cost, providing improved energy efficiency for servers.

LPC2148 is the widely used IC from ARM-7 family. It is manufactured by Philips and it is pre-loaded with many inbuilt peripherals making it more efficient and a reliable option for the beginners as well as high end application developer.

MAX-232 version of serial I/O standard is most widely used in PCs, GSM/GPS and several devices. In MAX232, high and low bits are represented by fallowing voltage ranges. The MAX232 has 2 sets of line drivers for transferring and receiving data's shown in fig. The line drivers used for TxD are called T1 and T2, while the line drivers RxD are designated as R1 and R2. The MAX232 is an integrated circuit that converts signals from an RS-232 serial port to signals suitable for use in TTL compatible digital logic circuits. The MAX232 is a dual driver/receiver and typically converts the RX, TX, CTS and RTS signals.

The drivers provide RS-232 voltage level outputs (approx. ± 7.5 V) from a single + 5 V supply via on-chip charge pumps and external capacitors. This makes it useful for implementing RS-232 in devices that otherwise do not need any voltages outside the 0 V to + 5 V range, as power supply design does not need to be made more complicated just for driving the RS-232 in this case. The receivers reduce RS-232 inputs (which may be as high as ± 25 V), to standard 5 V TTL levels. These receivers have a typical threshold of 1.3 V, and a typical hysteresis of 0.5 V. The later MAX232A is backwards compatible with the original MAX232 but may operate at higher baud rates and can use smaller external capacitors – 0.1μF in place of the 1.0μF capacitors used with the original device.

A GSM modem **is** a specialized type of modem which accepts a SIM card, and operates over a subscription to a mobile operator, just like a mobile phone. From the mobile operator perspective, a GSM modem looks just like a mobile phone. Designed for global market, SIM 300 is a Tri-band GSM/GPRS engine that work on frequency EGSM (Extended GSM) 900 MHZ, DCS(Digital cellular service) 1800 MHZ and PCS (Personal Communication Services) 1900 MHZ . SIM 300 provides GPRS multi-slot class 10 capability and support the GPRS (General Packet Radio Service) coding schemes CS-1, CS-2, CS-3 and CS-4.with a tiny configuration of 40mm x 33mm x 2.85mm, SIM 300 can fit almost all the space requirement in your application, such as smart phone, PDA, phone and other mobile device. The physical interface to the mobile application is made through a 60 pins board-to-board connector, which provides all hardware interface between the module and customer's boards expect the RF antenna interface.

AT commands are instructions used to control a modem. AT is the abbreviation of Attention. Every command line starts with 'AT' or 'at'. That is why modem commands are called AT commands. Many commands that are used to control wired dial up modems, such as ATD(Dial), ATA(answer), ATH(hook control) and ATO(return to online data state), are also supported by GSM/GPRS mobile phones. Besides this common AT commands set GSM/GPRS mobile phones support an AT command set that is specific to the GSM technology, which includes SMS related commands, like AT+CGMS (send message), AT+CMSS (send message from the storage), AT+CMGL (list messages) and AT+CMGR (read messages).Note that the starting "AT" prefix that informs the modem about the start of a command line. It is not the part of the AT command name. For example D is the actual AT command name in ATD and +CMGS is the actual AT command name in AT+CMGS, however some books and websites use them interchangeably as the name of an AT command.

A Liquid Crystal Display (LCD) is a thin, flat display device made up of any number of colors or monochrome pixels arrayed in front of a light source or reflector. It is preferred by engineers because it uses very small amount of electric power. Liquid crystal display (LCD) offers several advantages over traditional cathode ray tube that makes them ideal for several applications. Of course LCD's are flat and they use only a fraction of power required by cathode ray tubes. They are easier to read and more pleasant to work with for long periods of time. There are several tradeoffs as well, such as limited view angle, brightness and contrast, not to maintain

high manufacturing cost. 16x2 LCD is used in this project to display data to user. There are two rows and sixteen columns. It is possible to display 16 characters on each of the 2 rows. It has registers, command and data register.

Float Sensor is an electrical ON/OFF Switch, which operates automatically when liquid level goes up or down with respect to specified level. The Signal thus available from the Float Sensor can be utilized for control of a Motor Pump or an allied electrical element like Solenoid, Lamps, and Relays etc. Float Sensors contain hermetical sealed Reed Switch in the stem and a permanent Magnet in the Float. As the Float rises or falls with the level of liquid the Reed Switch is activated by Magnet in the Float.

You can measure temperature more accurately than a using a thermistor. The sensor circuitry is sealed and not subject to oxidation, etc. The LM35 generates a higher output voltage than thermocouples and may not require that the output voltage be amplified.

This is a water pump and valve where both are in a single component. A pump is driven using a voltage of 12V. This is interfaced to the arm controller indirectly using relay. This relay triggers the 12v voltage and interfaces with arm controller. Then the pump is driven and water is supplied using valve present in that component.

Relay is an electrically operated switch. It is used for voltage conversion from 12V to 5V. It consists of normal open, normal close and common pins on one side and on another side it has ground, enable and 12V where input is given to arm controller and Normal open and common are shorted and given to one end of pump.

Soil moisture sensors measure the water content in soil. A soil moisture probe is made up of multiple soil moisture sensors. This Soil Moisture Sensor can be used to detect the moisture of soil or judge if there is water around the sensor, let the plants in your garden reach out for human help. Insert this module into the soil and then adjust the on-board potentiometer to adjust the sensitivity. The sensor would outputs logic HIGH/LOW when the moisture is higher/lower than the threshold set by the potentiometer. With help of this sensor, it will be realizable to make the plant remind you: Hey, I am thirsty now, please give me some water.

## V. WORKING

The working of the system is as follows: ARM microcontroller is interfaced with soil moisture sensor, water level sensor, GSM, LCD, Temperature sensor and water pump. First initialize GSM and wait until it obtains the network. The green light indicates GSM is ON and red light indicates the network strength. Once it obtains complete network there will be delay in blinking of red light. Then initialize LCD, it will display all the statements given in code such as title, and the status of field regarding pump and tank storage.

Using the inputs from soil moisture sensor and water level sensor we turn ON the water pump manually as well as automatically. When the soil is moisturized and soil moisture sensor senses moisture presence it keeps the pump off and when it senses absence of moisture it switches on the pump and supply water to field. Then we need to continuously monitors the sensors and based on that we take some actions. It supplies water until it again sense the presence of moisture, once it senses moisture is present is switches of the pump. Water level sensor gives the presence of water in tank whether it is high or low. Temperature sensor gives room temperature. This whole process is automatic process of irrigation system using wireless sensor networks and GSM module.

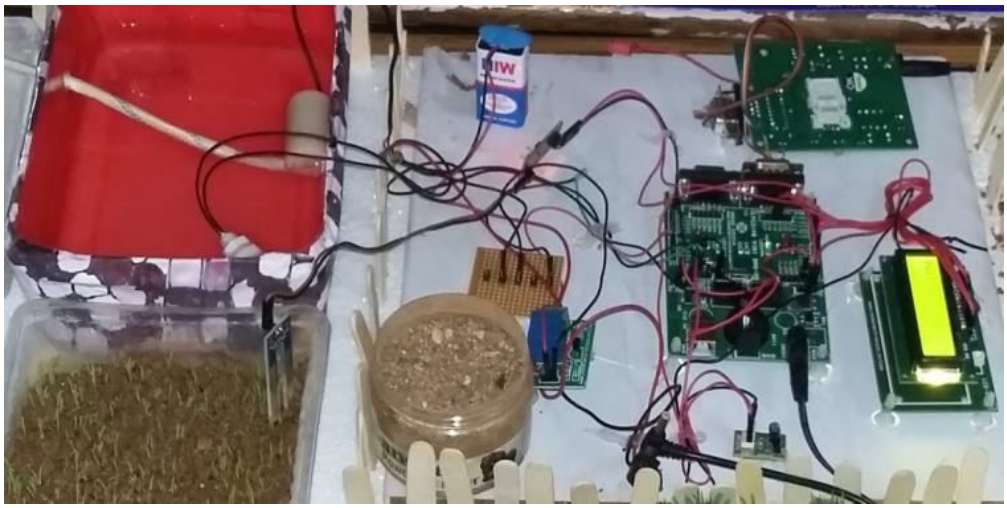


**Fig.2 Irrigation System**

Now we look at the software application part of the irrigation system which is a manual process.


**Fig.3 Software Application**

This software application is written using JAVA code which has 2 main options pump on/off and status. This is done using GSM which can be used from any place framer is. Then he clicks on pump off it sends a message and informs to switch off the pump which will then intimate by sending a SMS that pump is switched off. Then farmer clicks on pump ON here 2 activities take place: If the field is wet and when farmer switches on the pump, it will inform to switch on and intimates by sending a SMS that pump is ON but it senses that field is wet and switches the pump off and intimates pump off by sending SMS. If the field is wet and when farmer clicks on switch ON pump it will inform to switch ON pump and intimates by sending a SMS that pump is on. If farmer wants to know the status of field whether the soil is dry or wet and tank storage is low or high, farmer can get to know these details by clicking on status. Farmer will receive a message as tank is empty/full and water storage is high/low. There is another method controlling field: If a farmer is not having an android mobile and farmer has no software application, a farmer can use normal mobile by sending normal text messages as follows:

$1- indicates pump ON.

$2- indicates pump off.

$3- indicates status.

**ALGORITHM**

1.) INITILAISE GSM:

unsigned char cmd_1[]="AT" //attention command

```
unsigned char cmd_2[]="ATE0"; //charecters not echoed
unsigned char cmd_3[]="AT+CNMI=2,1,0,0,0"; //procedure for message reception from the n/w
unsigned char cmd_4[]="AT+CMGF=1"; //set text mode
unsigned char cmd_6[]="AT&W"; //modification saving in eeprom
```

2.) UART INITIALISE:

```
PINSEL0|=0X000000005;
U0LCR=0X83;
U0DLL=0X61;
U0DLM=0X00;
U0LCR=0X03
```

3.) ARM:

```
 PINSEL1|=0X15400000;
ADCR =0X00200404; // FRM MSB- SCBITS=1,OPNL MODE, 8BITMOD,CLK FREQ,4th CHANNEL
ADCR|=0X01000000;
while(!(ADDR & 0X80000000));  //WHEN DONE BIT IS NOT HIGH BE HERe val=ADDR;
val>>=6;
val&=0x000003ff;
val=val/5;
return(val);
```

4.) IODIR1 = 0X00200000;  // FOR WATER PUMP

```
IODIR0 &=0x00100000;  //// for water float
if(water_float1==0X00100000)  //// for water float
IODIR0 & = 0x00000400; ////for moisture  if(moisture==0X00000000)  ////p0.10 for moisture
```

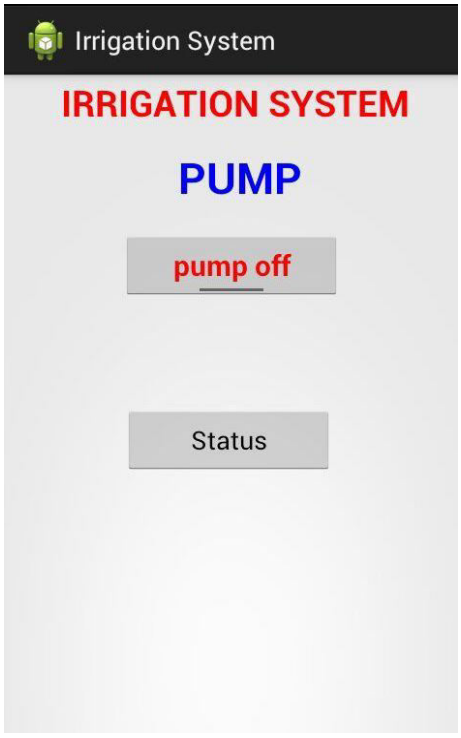



**Fig 4 Android Application**

The simulation part is as follows:

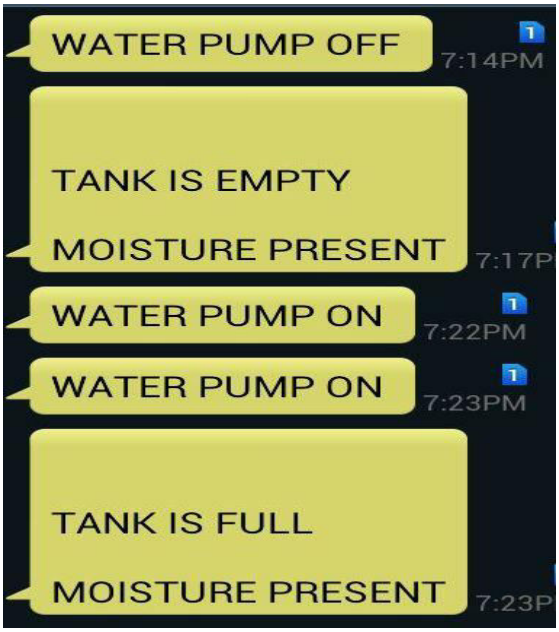



**Fig 5 Simulation User Receives Messages**

## VI. ADVANTAGES

The system is very economical in terms of hardware component and power consumption. The system helps in saving of water and electricity. It can be implemented in large agricultural areas. With the help of GSM user can control the motor from anywhere by just sending SMS. The system helps in labor problem when there are no labors to work and eliminates man power. System can be switched into manual mode whenever required. It is useful to all climatic conditions and all types of irrigation.

## VII. APPLICATIONS

Irrigation can be done in fields, gardens, farms etc. It is efficient for varieties of crops. This implementation can be used for patient monitoring. The software application developed for this system can be used for household works such as tank storage. This system can be operated automatically as well as manually.

## VIII. FUTURE SCOPE

Rain gun sensor can be added so that when it rains there won't be floods and this shield the field and avoids floods. Rain water harvesting can be done and this harvested water can be used to irrigate fields. Hooters can be used so that it gives siren at various occasions such as intrusion detection, floods etc. Using IR sensors any object passing into fields can be detected and alerted.

## IX. CONCLUSION

The automated irrigation system implemented is very feasible and cost effective. The system is very economical in terms of hardware component and power consumption. The system helps in saving of water and electricity. It can be implemented in large agricultural areas. With the help of GSM user can control the motor from anywhere

by just sending SMS. The system helps in labor problem when there are no labors to work and eliminates man power. System can be switched into manual mode whenever required. It is very useful in all climatic conditions and all types of irrigation. In this solar power can be made use of where cost of electric power would be expensive. It can be implemented for variety of crops.

## X. ACKNOWLEDGEMENT

The successful completion of project would be incomplete without complementing those who made it possible and whose guidance and encouragement made our efforts successful. With deep sense of gratitude we acknowledge the help and encouragement of our project guide, Assoc. Prof. **Mrs Sujatha S.**, CMRIT Bangalore, for her successful guidance, support, help and suggestions. We are thankful to our Head of Department, Mrs Sharmila K.P. and Senior Professor Dr. K. V. S. Ananda Babu, CMRIT for their support and enquiry towards our project. We also take this opportunity to extend our regards to Dr. Sanjay Chitnis Principal, CMRIT for his support. We would also like to thanks all those who were involved directly or indirectly and who helped in our project. Finally, we would like to thanks our parents and friends who have bestowed faith in our ability and have helped us stand in a place where we are.

## REFERENCES

[1] Veena Divyak ,Ayush Akhouri,A Real time implementation of a GSM based Automated Irrigation Control System using drip Irrigation Methology (Volume 4, Issue 5,May 2013)

[2] Mansour ,H.A, You sifEl-Melhem **,**impact tthe automatic control of closed circuits raingun irrigation system on yellow corn growth and yield(International Journal of Advanced Research (2013), Volume 1, Issue 10, 33-42)

[3] m. guerbaoui, y. el afou, a. ed-dahhak, a. lachhabpc-based automated drip irrigation system(Vol. 5 No.01 January 2013)

[4] Purnima, S.R.N Reddy, "Design of Remote Monitoring and Control System with Automatic Irrigation System using GSM Bluetooth", on IJCA,2012

[5] Cottef F. (2001): LabVIEW: programmationet applications. Dunod, pp. 415.

[6] Eddahhak, A.; Lachhab, A.; Ezzine, L.; Bouchikhi, B. (2007): Performance evaluation of a developing greenhouse climate control with a computer system. AMSE Journal Modelling C, 68 (1), pp. 53-64.

[7] Elattir H. (2005): La conduiteet le pilotage de l'irrigation goutte-àgoutte en maraîchage. Bulletin mensueld' informationet de liaison duprogramme national de transfert de technologie en agriculture (PNTTA), pp. 124

[8] Gonzalez, R .A.; Struve, D.K.; Brown L.C. (1992): A computer controlled raingun irrigation system for container plant production Hort Technology, 2(3), pp. 402-407.

[9] Howell, T.A. (2001): Enhancing water use efficiency in irrigated agriculture. Agron. J, 93, pp. 281-289.

[10] AbouKheira A. A. **(2009).** Comparison among different irrigation systems for deficit- irrigated transgenic and non transgenic yellow corn in the Nile Valley". Agricultural Engineering International: the CIGR Ejournal. Manuscript LW 08 010. Vol. XI:1 25.

[11] "Philosophy and Goals". Android Open Source Project.Google. Retrieved2012-04-21.

[12] P. Varady, Z. Benyo, and B. Benyo, "An open architecture patient monitoring system using standard technologies," IEEE Trans. Inf. Technol. Biomed., vol. 6, no. 1, pp. 95–98, Mar. 2002.

[13] J. Bai et al., "The design and preliminary evaluation of a home electrocardiography and blood pressure monitoring network," J. Telmed. Telecare, vol. 2, no. 2, pp. 100-06, 1996.

[14] G. Williams, P. J. King, A. M. Capper, and K. Doughty, "The electronic doctor (TED)—A home telecare system," in Proc. 18th IEEE Annu. EMBS Int. Conf., Amsterdam, The Netherlands, Oct. 31-ov. 3, 1996, vol. 1, pp. 53-4.

[15] P. Johnson and D. C. Andrews, "Remote continuous physiological monitoring in the home,"

[16] J. Telmed. Telecare, vol. 2, no. 2, pp. 107-13, 1996. K. Doughty, K. Cameron, and P. Garner, "Three generations of telecare of elderly," J. Telmed. Telecare, vol. 2, no. 2, pp. 71-0, 1996.

[17] T. Fukatsu and M. Hirafuji, "Field monitoring using sensor-nodes with a web server," J. of Robotics and Mechatronics, vol. 17, no. 2, pp. 164-172, 2005.

[18] K. Langendoen, A. Baggio, and O.Visser,"Murphy loves potatoes: experiences from a pilot sensor network deployment in precision agriculture," The 20th Int. Parallel and Distributed Processing Symposium, pp. 25-29, April 2006.

[19] N. Wang, N. Zhang, M. Wang, "Wireless sensors in agriculture and food industry—recent development and future perspective," Comp. Electron. Agric. vol. 50, no. 1, pp. 1-14 2006.

[20] T. Fukatsu and M. Hirafuji, "Field monitoring using sensor-nodes with a web server," J. of Robotics and Mechatronics, vol. 17, no. 2, pp. 164-172, 2005.

[21] Sunnorth, SPCE061A compiler user menu v1.0, www.sunnorth.com.cn. Spectrum Technologies Inc., External temperature sensor category #3667.

[22] E. P. Eldredge, C. C. Shock, and T. D. Stieber, "Calibration of granular matrix sensors for irrigation management," Agronomy Journal, vol. 85, pp.1228-1232, 1993.

[23] S. J. Thomson, T. Youmos, and K. Wood, "Evaluation of calibration equations and application methods for the Watermark granular matrix soil moisture sensor," Appl. Eng. Agric., vol. 12, pp. 99-103, 1996.

[24] Electronique-Diffusion Company Inc., REHS135, www.elecdif.com.

[25] Allguy International Co., Ltd., CDS photoconductive cells, http://cn.commerce.com.tw/modules.php?modules=company&action=company_inside&ID=A0001712&s=h.Ready International Inc., ZigBee 3160 module, http://www.ritii.com/ch/

[26] M. Yu, H. Mokhtar, and M. Merabti, "A survey on Management in wireless sensor networks," www.cms.livjm.ac.uk/pgnet2007/Proceedings/Papers /2007-099.pdf. www.intechopen.com

[27] Wireless Sensor Networks and Their Applications to the Healthcare and Precision Agriculture 21 Sapon Tanachaiwiwat, Pinalkumar Dave, Rohan Bhindwale, Ahmed Helmy, "Secure locations: routing on trust and isolating compromised sensors in location-aware sensor networks," SenSys'03, pp. 324-325, LA, CA, USA, Nov. 5–7, 2003.

[28] S Harte, A Rahman, K M Razeeb, "Fault tolerance in sensor networks using self-diagnosing sensor Node," The IEE Int. Workshop on Intelligent Environments, pp. 7-12, 2005.

[29] D. Estrin, R. Govindan, J. S. Heidemann, S. Kumar, "Next century challenges: scalable coordination in sensor networks," In Mobile Computing and Networking, pp. 263-270, 1999.

[30] A. T. Tai, K. S. Tso, W. H. Sanders. "Cluster-based failure detection service for large-scale ad hoc wireless network applications in dependable systems and networks,"

[31] Proceedings of the 2004 Int. Conf. on Dependable Systems and Networks, DSN ' 04. 2004.

[32] T. Clouqueur, K. Saluja, P. Ramanathan, "Fault tolerance in collaborative sensor networks for target detection," IEEE Transactions on Computers, vol. 53, pp. 320-333, 2004.

[33] M. Ding, D. Chen, K. Xing, X. Cheng. "Localized fault- tolerant event boundary detection in sensor networks," in Proceedings of INFOCOM, 2005.

[34] J. Chen, S. Kher, and A. Somani, "Distributed fault detection of wireless sensor networks," DIWANS'06, Los Angeles, USA:ACM Pres, pp. 65-72, 2006.

# STUDY OF MICROSTRIP TAPPED HAIRPIN RESONATOR FILTERS: A REVIEW

**Alpana Singh[1], Tejbir Singh[2]**

[1]*M.Tech Scholar, Subharti University, Meerut, U.P.(India)*

[2]*Assistant Professor, EC Department Subharti University, Meerut, U.P. (India)*

## ABSTRACT

*Practical design techniques are presented for tapped hairpin resonator filters on FR4 laminates. The hairpin filter is one of the most popular low microwave frequency filters because of it is compact and does not require grounding. Its design on FR4 laminates is very difficult to do because of the relatively poor performance of the laminate at the microwave region. The laminate properties of the FR4 become nonlinear unlike more expensive microwave laminates. The motivation to use FR4 in the low microwave frequencies is its cost. Methods and techniques were developed to address these problems of the FR4 and are discussed in this paper.*

***Keywords: Bandpass, Hairpin Filter***

## I. INTRODUCTION

The most popular microstrip filter in low microwave frequencies is the hairpin resonator filter. It does not require a large real estate as the edge-coupled filter. It also does not require critical grounding thus making manufacturing easier. There is however, very little literature that discusses the design of this filter. CST MWS and HFSS is normally employed to synthesize and optimize the design for the desired specifications and yield. A very important requirement for the HFSS simulation results to be close to actual results is correct modeling of the filter structure and precise characterization of the substrate or laminate used.

The FR4 laminate is the most common electronic carrier for circuits operating below the microwave frequency. Beyond 1 GHz, the laminate properties (like $\varepsilon_r$, tan $\delta$, roughness, etc…) of the FR4 become nonlinear. It becomes difficult to precisely characterize the FR4 at the desired microwave frequency. Beyond 3 GHz, the use of the FR4 is not anymore recommended because of unacceptable attenuation. (There are however some amateur designs [1] that employ FR4 up to 5.7 GHz but these modules are very lossy for industry standards.) Choosing values close to the data sheet values and values derived from analyzing quarter and half wavelength microstrip lines normally result in HFSS simulation results that are off the actual response by several hundred megahertz. The FR4 however is cheaper than most stable microwave laminates that methods were developed so that filters can be designed without the aid of simulation and synthesis software.

## II. INITIAL HAIRPIN FILTER DESIGN

The hairpin filter is a variant of the edge-coupled band pass filter. A sliding factor is introduced to allow for bending thus making the design more compact (see Fig. 1) [2], [3]. The details of designing an edge-coupled filter are described in [4].

The structure depicted in Fig. 1c however, was observed to introduce a large insertion loss and an unacceptable return loss in FR4 laminates. A tapped hairpin resonator filter has a smaller insertion loss and better return loss. This is a variant of an edge-coupled filter that contains a matching stub (see Fig. 2). Varying the length of this stub and the tapping distance varies the return loss and consequently the insertion loss.

Optimizing the design using HFSS software where the substrate parameters are:

dielectric constant, $\varepsilon_r = 4.77$

laminate height, $H = 59$ *mil*

metal thickness, $T = 0.7$ *mil*

loss Tangent, Tan $\delta = 0.008$

roughness = 0.075 mil

relative permeability, $Mur = 1$

conductor conductivity, $Cond = 5.88E+07$

*cover height = 3.9E+34 mil*

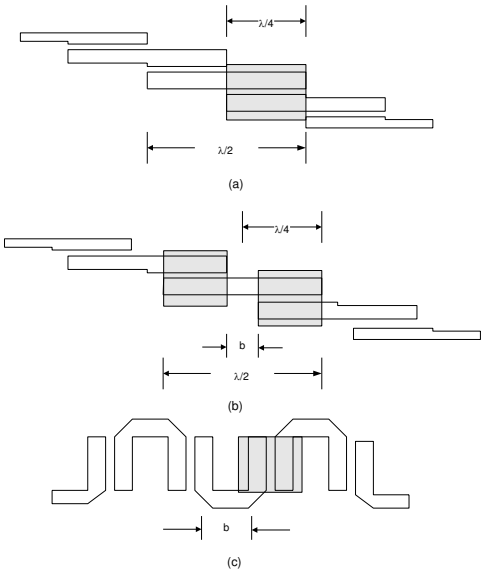



**Figure 1. (a) Edge-Coupled Filter (b) The Resonators are Moved to Provide for a Slide Factor, *b***

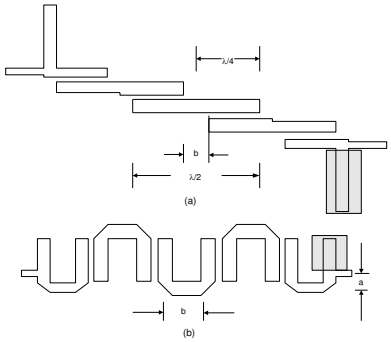



**Figure 2. (A) Edge-Coupled Filter with Matching Stub (B) The Matching Stub is Incorporated into the Resonator Thereby Forming the Tapped Hairpin Resonator Filter**

often lead to results that are several hundreds of megahertz away from the actual results [5]. Patterns on the behavior of the filter were observed during the characterization and testing of different filter samples. Each

sample has a specific structural dimension altered to test the effect on the filter response. These patterns are presented in the next section.

## III. DESIGN PATTERNS

The following design patterns were observed while a filter is being designed experimentally for the Digital Microwave Radio project of the Advanced Science and Technology Institute. The filter must be able to operate between 2.32-2.36 GHz. The filter response at 2.45 GHz should be about 15 dB below the average insertion loss between 2.32-2.36 GHz thereby providing an additional 15 dB LO rejection to the mixer. The insertion loss at 2.54-2.58 GHz should be as low as possible. A second channel requires a filter with a passband at 2.54-2.58 GHz but rejects the 2.32-2.36 GHz frequencies.

The filters made throughout the experiment are $5^{th}$ order filter. Fewer orders would provide poor selectivity and more orders would make the design too large.

### 3.1 Initial Design

An edge-coupled filter is first designed using the steps enumerated in [4]. The $Z_{oo}$, $Z_{oe}$ values and the laminate parameters are then sent to the LineCalc transmission line calculator [6] to determine the length of the resonator, the spacing and the width. A slide factor is then introduced to provide for the bending of the resonator into a hairpin structure. The basic structure is shown in Fig. 3 including the dimensions varied to determine the specific response of the filter.

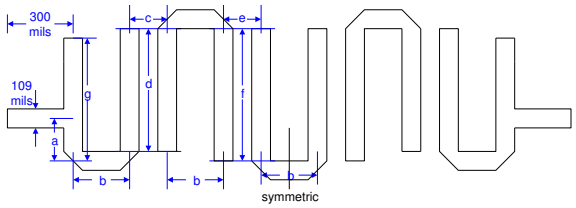



**Figure 3. Tapped Hairpin Resonator Filter Dimensions. The Structure is Assumed to be Symmetric So that $S_{11}=S_{22}$ and $S_{21}=S_{22}$ and that Fewer Variables Could be Used**

The dimensions used in most of the experiments were:

*a = 263 mils*

*b = 525 mils*

*c = 131 mils*

*d = 416 mils*

*e = 171 mils*

*f = 416 mils*

*g = 430 mils (1 mil*= 0.00254 centimeters*)*

The width of the $1^{st}$ resonator according to the steps given in [4] should be 91 mils while the $2^{nd}$ and $3^{rd}$ resonators should have widths of 107.6 and 108 mils. For simplicity of design, the width of the structure within the filter was set at 108 mils while the tapping line is 109 mils (50 ohm).

### 3.2 Coupling Resonator Length, *d*

Holding every variable in Fig. 3 constant except for d, designs were made for *a = 263 mils, a + 2 mm, a + 4 mm, a – 2 mm* and *a – 4 mm*. The center frequencies were noted and plotted in Fig. 4. This curve can be a useful reference when designing for the center frequency of the hairpin filter. Note however that this curve is relative to the initial values of parameters *a* to *g*. This means that if we want frequency *x* and this corresponds to coupling resonator length *y* and when the results of the actual test is off the target frequency *x* by $\bar{}x$ (because the parameters *a* to *g* are not the same), all we have to do is estimate from the curve the $\bar{}y$ needed to arrive at the desired frequency. Naturally, if the resonator length has been set to the desired frequency and other variables were changed, the center frequency may again change but this is not as severe as the effect of the resonator lengths on the center frequency. These changes can easily be compensated in succeeding design iterations.

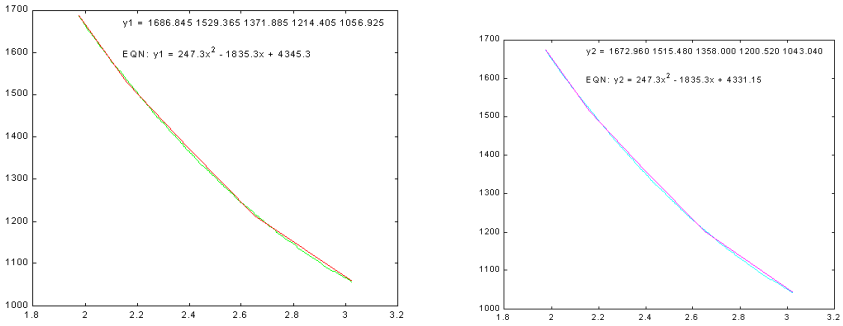



**Figure 4. (a) Curve for the 1st Resonator Length. The y-axis Gives the Total Length (i.e. *g + b + d*). (b) Curve for the 2st and 3rd Resonator Lengths. The y-axis Gives the Total Length (i.e. *d + b + f* or *2f + b*).**

The equations of the of the curves in Fig. 4 are:

1st resonator length $y = 247.3\ f_c^2 - 1835.3\ f_c + 4345.3$ (1)

2nd,3rd resonator length $y = 247.3\ f_c^2 - 1835.3\ f_c + 4331.15$ (2)

### 3.3 Resonator Spacing, *c* and *e*

The rule of thumb used in this design is *e = c + 1 mm* since more coupling is required in the input matching/resonating element. This assumption allows the manipulation of just one variable instead of two thereby making analysis easier. A more involved experiment can be done to determine the actual response of varying *c* and *e*.

Smaller resonator spacing provide more coupling between adjacent resonators. More signal power is transferred and the resulting filter response is a wider bandwidth as seen in Fig. 5. Ripples are also present of the bandwidth is too wide.

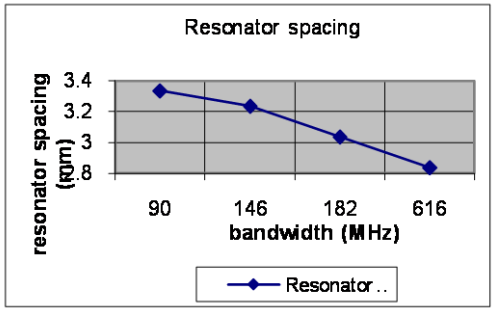


**Figure 5. Resonator Spacing vs. Bandwidth. The Desired Bandwidth Can be Initially Estimated from this Curve.**

### 3.4 Sliding Factor, *b*

The sliding factor *b* affects the insertion loss, the center frequency and the possibility of reentrant frequencies. Decreasing *b* lowers the insertion loss, as more signal power is concentrated on the coupling structures instead of being attenuated along a transmission line path. However, if *b* is decreased without holding the total resonator length (given in Fig. 4) constant, then a frequency shift occurs. Also, decreasing *b* drastically may introduce resonator self-coupling, which produces reentrant frequencies [3].

Experiments on decreasing b while keeping all other variables constant were done and the results are frequency shift and varying $S_{11}$ response.

Experiments on decreasing *b* while keeping the total resonator length constant were also done and the results are slight changes in bandwidth and varying S11 response. This draws a conclusion that by keeping the total resonator length constant, the $S_{11}$ can be controlled. The effect of the sliding factor on the $S_{11}$ is more significantly observed when the 1st set of sliding factor experiments was performed. It is therefore recommended to vary only the sliding factor *b* to scan for the best $S_{11}$ response. Adjustments in resonator length can later be done to obtain again the desired center frequency.

### 3.5 Tap Distance, *a*

The tap distance *a* controls the return loss. Since the tapped hairpin resonator filter is actually a modified edge-coupled filter with matching stubs (see Fig. 2), adjusting *a* is just like adjusting the tuning stub's length to obtain an impedance match. Like most variables, there is an optimal value below and beyond which the response deteriorates. It was observed that other properties of the filter like bandwidth and center frequency are almost not affected when *a* is varied. It was also observed that the insertion loss of the filter decreases as the return loss is increased. The tap distance is therefore the last variable being varied for optimal design.

## IV. PRACTICAL TECHNIQUES

To design a hairpin filter, first follow the steps outlined in Section 3.1 of this paper to obtain the initial values of the variables. A prototype filter is then done and tested. The center frequency and bandwidth are then noted.

The center frequency results are compared with the curves of Fig. 4. The total resonator lengths are then adjusted following the explanation given in Sec. 3.2. Using the same technique, the resonator spacing is deduced from the curve of Fig. 5 for the desired bandwidth. Create the next prototype and then test to verify. The resulting center frequency and bandwidth may not be exactly as planned but they should be close the desired values.

In choosing the right sliding factor, the practical technique is to first use an arbitrary sliding factor (*b = 525 mils* in the case of the experiment). Choose the total resonator length for a particular frequency based on Fig. 4. Vary *b* by a constant amount (*b = 525 mils – 2n mm* where n = 1, 2, 3, … in the case of the experiment) while holding *d*, *f* and *g* constant. The center frequency will surely shift. Choose the response with the best return loss and the least insertion loss. (This part of the design iteration may take 3 prototype boards) This is approximately the correct sliding factor for the filter. Readjust *d*, *f* and *g* using Fig. 4 as basis to obtain the desired center frequency. The change in the insertion loss and return loss will just be slight.

Once the optimal resonator lengths, resonator spacing and sliding factor have been chosen, design a few more boards with varying tap distance to get the optimal design. Do not go beyond half of the resonator length *g* as the response would have been terrible by then. By rule of thumb, the bandwidth should have a return loss of 10 dB at least.

The total number of boards to be used in this design should be about 8-10 boards. This is a lot less than the 32 boards used in this research to produce an optimal 2.32-2.36 GHz band pass filter. The final optimal filter produced is shown in Fig. 6 and its results in Fig. 7. The final filter has a 6 dB insertion loss, a very good result considering the fact that small and steep bandwidths generally suffer from high insertion loss and that above 2 GHz, the FR4 laminate is very lossy.

## V. CONCLUSION

Practical techniques based on design patterns of a tapped hairpin resonator filter on FR4 laminates have been presented in this paper. This motivation is due to the difficulty of designing such filters on FR4 laminates using HFSS. Several patterns have been observed. The resonator length was shown to have a significant effect on the center frequency. The resonator spacing generally controls the bandwidth. Choosing the right slide factor increases the potential of the filter for a good match and low loss. The tap distance ultimately matches the filter to the rest of the circuitry. A step-by-step method is then presented in order to aid in the filter design.

## REFERENCES

[1] ARRL, Ed. The ARRL UHF/Microwave Projects Manual. CT: ARRL, 1994.

[2] E. Cristal and S. Frankel, “Hairpin-Line and Hybrid Hairpin-Line/Half-Wave Parallel-Coupled-Line Filters,” IEEE Trans. on MTT, vol. 20, no. 11, pp. 719-728, November 1972.

[3] R. Rhea, HF Filter Design and Computer Simulation. Atlanta, GA: Noble Publishing, 1994.

[4] G. Matthaei, L. Young, and E. M. T. Jones, Microwave Filters, Impedance-Matching Networks, and Coupling Structures. Boston, MA: Artech House, 1980.

[5] T. Laverghetta, Microwave Materials and Fabrication Techniques, 3rd Ed., Norwood, MA: Artech House, 2000.

[6] Agilent Technologies, “Line Calc Program,” Advanced Design System ver. 1.3, Palo Alto, CA, November 1999.

# FREQUENCY RECONFIGURABLE MICROSTRIP SLOT ANTENNA LOADED WITH VARACTOR DIODE

**RitikaTandon[1], Tejbir Singh[2]**

[1]*M.Tech Scholar, Subharti University, Meerut, U.P. (India)*

[2]*Assistant Professor, EC Department, Subharti University, Meerut, U.P. (India)*

## ABSTRACT

*This paper presents the simulated results of a compact varactor diode integrated conventional circular patch antenna and also compares the results with a similar patch antenna. Proposed microstrip patch antenna (MSA) was designed and simulated using HFSS.V.13 and CST MWS and its various parameters such as return loss, VSWR and input impedance were determined, and shape of this MSA was modified by cutting various slots in it at appropriate positions. The diode is modeled as a switch for the frequency band from 2.85 to 3.0 GHz. The proposed antenna also gives CP radiation with slots. In addition to it is observed that the proposed antenna shows frequency agility behavior in the frequency ranges 2.85 to 3.0 GHz with bias voltage varying from 0 to 5V or in ON and OFF state. A bandwidth enhancement and miniaturization is also achieved.*

***Keywords: Bandwidth enhancement; Circular polarization(CP); Varactor diode***

## I. INTRODUCTION

Microstrip patch antenna consists of a metallic radiating patch backed up by a dielectric substrate and a ground plane below that. Now days, MSAs is widely used in many applications due to their advantages such as low profile, lightweight, planer configuration and ease of fabrication. However the main limitation of MSAs is their inherently narrow BW [1]. In order to improve the BW of MSA, the idea of integrating active devices has been implemented for last many years. Such types of antennas are known as active integrated antenna. The active integrated antenna (AIA) has been a growing area of research in recent years.

An AIA can be regarded as an active microwave circuit in which the output or input port is free space instead of a conventional 50 Ω interface. Active antennas reduce size, weight and cost over conventional designs which are very useful in microwave systems [2-4]. Active antennas overcome several limitations of traditional microstrip antennas [5]. They are almost frequency independent that is their bandwidth is depends on the active circuitry rather than the radiating element. Also a careful design of the connected amplifier may ensure broadband characteristics of the antenna. Similarly the gain of the antenna can also be controlled by using the amplifiers. Since the active antenna is electrically small compared to the passive one, the overall length is much less than the conventional antenna, and thus can be used in places where there is limitation of space.

Integrated Antennas and Active integrated antennas (AIAs) are widely used in the area of wireless communications, both for civilian and military purposes. In particular, AIAs are devices in which a passive antenna element and an active circuitry are integrated together on the same substrate. The integration of active solid state devices like oscillators, varactor diode, gun diode, amplifiers,and mixers grants greater compactness, lower costs and higher power efficiencies with respect to conventional passive layouts[6-11].

In this paper, a varactor diode has been modeled in CST MWS and integrated with a hexagonal microstrip patch antenna with unequal side to form an active antenna latter which is modeled in HFSS and CST. The reason behind selecting the hexagonal microstrip antenna that, it has smaller size compared to the square and circular microstrip antennas, as well as better impedance bandwidth over rectangular and square microstrip antennas for a given frequency. Therefore, authors have designed a coaxial fed hexagonal patch antenna and circularly polarized radiation has been achieved by adjusting the position across the antenna. It is found that the obtained results are encouraging for practical applications. As it enhances BW of HMSA a comparative analysis of the various geometries of MSA obtained by cutting slots inside the radiating patch indicate considerable improvement in BW without much sacrifice on other performance parameters of MSA such as return loss, VSWR and its input impedance. Also when slots are inside the radiating patch it shows the good circular polarization

Frequency agility is the ability of a radar system to quickly shift its operating frequency to account for atmospheric effects, jamming, mutual interference with friendly sources, or to make it more difficult to locate the radar broadcaster through radio direction finding [12]. Frequency agility behavior of the proposed antenna will also be used in these applications.

## II. CALCULATIONS

*(i )Resonant frequency*

The resonance frequency of a CMSA is obtained using the given formula [3].

$$f_o = \frac{K_{mn}c}{2\pi a_e\sqrt{\varepsilon_e}}$$ where $K_{nm}$ is the $m$th root of the derivative of the Bessel function of order $n$ . For the

fundamental $TM_{11}$ mode, the value of $K_{nm}$ is 1.84118. The $a_e$ and $\varepsilon_e$ are the effective radius and the effective dielectric constant of the MSA, respectively. The fringing fields along the circumference of the given MSA are taken into account by replacing the patch radius $a$ by the effective radius $a_e$.

$$a_e = a\left[1+\frac{2h}{\pi a\varepsilon_r}\left\{\ln\left(\frac{a}{2h}\right)+1.41\varepsilon_r+1.77+\frac{h}{a}(0.268\varepsilon_r+1.65)\right\}\right]^{\frac{1}{2}} \quad (6)$$

The value of $\varepsilon_e$ is obtained using

$$\varepsilon_e = C(a, h, \varepsilon_e, \varepsilon_r)/C(a, h, \varepsilon_e) \quad (7)$$

where $C(a, h, \varepsilon_e, \varepsilon_r)$ and $C(a, h, \varepsilon_e)$ are the total capacitances of the dominant $TM_{11}$ mode of MSA with and without a dielectric substrate respectively. These can be calculated as [3]

$$C(a.h,\varepsilon_o,\varepsilon_r) = \frac{0.8525\varepsilon_o\varepsilon_r\pi a^2}{h}+0.5C_f \quad (8)$$ In (8), the first term is the main capacitance of the disc and the second

term is the fringing capacitance, $C_f$, which is given by

$$C_f = 2a\varepsilon_o\left[\ln\left(\frac{a}{2h}\right)+1.41\varepsilon_r+1.77+\frac{h}{a}(0.268\varepsilon_r+1.65)\right] \quad (9)$$

$C(a, h, \varepsilon_e)$ is calculated by putting $\varepsilon_r = 1$ in (8) and (9). For thin substrates, $\varepsilon_r$ should be used instead of $\varepsilon_e$ in (6.1), and for thick substrates ($h > 0.05\lambda_o$), $\varepsilon_e$ should be used.

*(ii) Actual radius of the Patch*

Using equations 6 to 9 and taking the values of different parameters as follows, we can calculate the value of actual radius of the patch.

$K_{mn}$=1.84118; c=3x10$^8$m/s;

$\varepsilon_o$= 8.86x10$^{-12}$F/m; $\varepsilon_r$=4.2;

h=1.6 mm; Frequency($f_o$)=2.1GHz

We can now calculate the value of the fringing capacitance $C_f$, $C(a, h, \varepsilon_e, \varepsilon_r)$, $C(a,h, \varepsilon_e)$, $\varepsilon_e$ and effective radius $a_e$ followed by the value of the actual radius 'a' which come out to be 24mm.

## III. DESIGN OF CIRCULAR MSA

The geometrical configurations of top and side views of the circular patch antenna are shown in Fig 1 & Fig 2. The resonant frequency for the dominant as well as for the higher order modes can be calculated from the formula of the circular disc by replacing radius *a* by equivalent radius $a_{eq}$[6-7] i.e.

$$f_{np} = \frac{X'_{np}.c}{2\pi a_{eq}\sqrt{\varepsilon_r}} \qquad (1)$$

Where $X'_{np}$ are the zeros of the derivative of the Bessel function $J_n(x)$ of the order *n* and the equivalent radius $a_{eq}$ is determined by comparing areas of a regular hexagon and a circular disk of radius $a_{eq}$ i.e.

$$\pi a_{eq}^{\ 2} = \frac{3\sqrt{3}S^2}{2} \qquad (2)$$

$a_{eq}$=0.9094S

Thus the resonant frequency of a hexagonal element may be expressed as:

$$f_r = \frac{X'_{np}.c}{2\pi(0.9094S).\sqrt{\varepsilon_r}} = \frac{1.1X'_{np}.c}{2\pi S.\sqrt{\varepsilon_r}} \qquad (3)$$

And for the lowest order mode $TE_{11}$

$X'_{np}$=1.84118

Hence $f_r$ =2.8 GHz

## IV. DESIGN SPECIFICATIONS AND CALCULATIONS

Design specifications for the proposed patch antenna are as follows**.**

| Parameter | Value |
|---|---|
| Substrate | FR4 |
| Thickness | 1.6 |
| Radius of Patch | 24 mm |
| Slot | (1x1.5)x11 in x and y direction |
| Diode/SW position | (2,2,1.6) and (-2,-2,1.6) |

## V. RESULTS AND DISCUSSION

Using these design parameters and mathematical expressions, the proposed antenna has been designed and performances are examined using HFSS and CST MWS, and the obtained results are described in the following

sections.

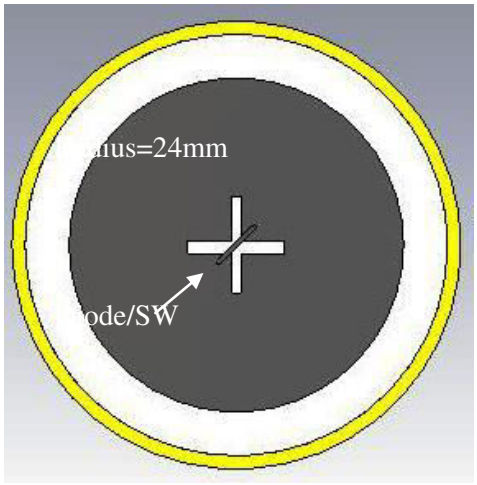



**Fig1. Top View of Reconfigurable Patch**

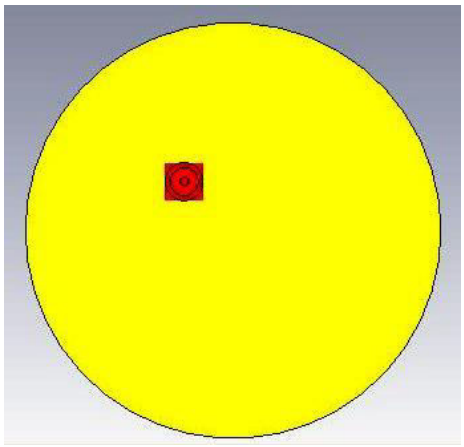


**Fig2. Bottom View**

Here a feed location point is to be found out on the conducting patch where patch impedance is 50 Ω. This feed point gives maximum radiation because of proper matching with 50 Ω coaxial feed. At first the feed position is varied and its effect on the input impedance, $S_{11}$ and VSWR are measured. The variation in return loss with frequency is shown in Fig 3,4 and 5 which shows that the varactor diode/Switch is in ON and OFF state.

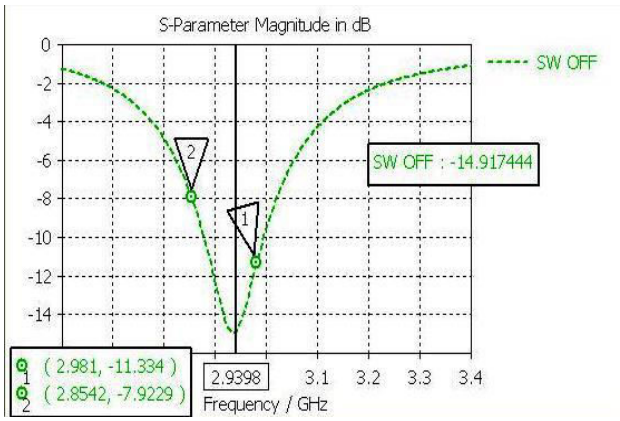



**Fig3. Return Loss when diode/SW is in OFF State**

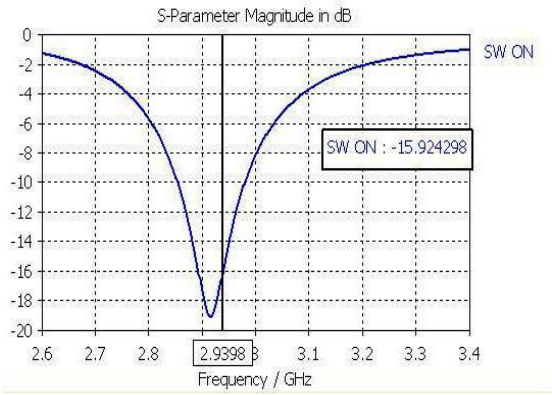



**Fig4. Return Loss when diode/SW is in ON State**

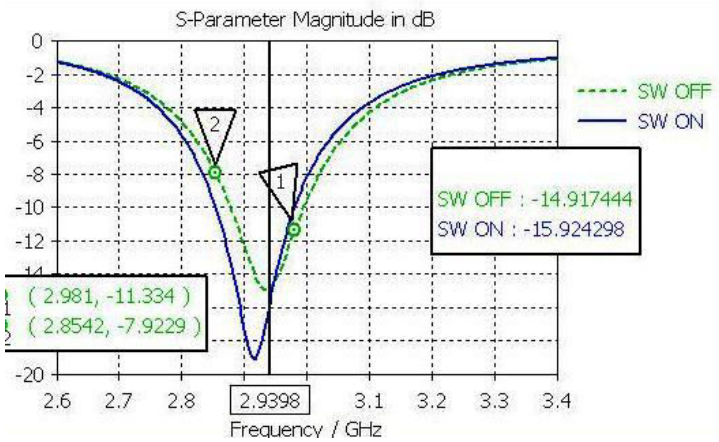



**Fig2. Comparative Return Loss when diode/SW is in OFF/ON State**

The radiation pattern near the resonant band frequencies shows that with an increase in frequency, the radiation pattern varies and the cross polar level increases significantly to the extent that the radiation becomes maximum

along $\Phi = 0°$ at 2.7 GHz. The radiation pattern of the antenna also shows that it is Omni directional as well as circularly polarized with small levels of cross polarization. The other parameters are also found as given in

**Table**

| Characteristics | Diode/SW ON | Diode/SW OFF |
|---|---|---|
| Resonance Freq | 2.91GHz | 2.94 GHz |
| Axial ratio (dB) | 0.255 | 1.48 |
| BW(GHz) | 0.132 | 0.127 |
| Gain (dBi) | 3.7 | 2.1 |

## VI. CONCLUSION

In this paper, design, simulation of conventional and varactor integrated circular patch antenna has been presented. The active device and the patch antenna have been considered as a composite unit instead of taking them as independent units whereas in conventional wireless or other communication systems antenna and circuit they are being considered as separate element. The antenna has been designed at 2.45GHz (ISM Band) and excited using coaxial feeding techniques and its performance characteristics such as return loss, axial ratio, VSWR, input impedance and radiation pattern has been calculated. The antenna with slots gives circular polarization (AR < 3 dB) as compared to the antenna without slots in addition to gain and bandwidth improvement. The differential bandwidth of antenna system when SW/Diode is in OFF state is 0.127 GHz whereas it becomes 0.132 GHz when the SW/diode is in ON position. The resonance frequency is also shifted from 2.94 GHz to 2.91 GHz clearly showing the property of antenna miniaturization.

## REFERENCES

[1] S.K. Behera, “Novel tuned rectangular patch antenna as a load for phase power combining", Ph.D. Thesis, 2007.

[2] C. M. Montiel, F. Lu and K. Chang, “Active-notch antennas stabilized with a slot line ring resonator for wireless applications,” IEEE Trans. on Antennas and Propagation, Vol. 46, No. 6, 945{946, 1998}.

[3] J. A. Navarro and K. Chang, “Integrated Active Antennas and Spatial Power Combining,” Wiley, New York, 1996, 1120 PIERS Proceedings, Kuala Lumpur, MALAYSIA, March 27-30, 2012.

[4] C. M. Montiel, L. Fan, and K. Chang, “A novel active antenna with self-mixing and wide-band varactor-tuning capabilities for communications and vehicle identification applications," IEEE Trans. Microwave Theory Tech., Vol. 44, pp.2421-2430, Dec. 1996.

[5] S. Sodagam, “Active receive antennas as sensors for radars", Technical Report CRSIS TR 107, Aug. 17, 2006

[6] R.K.Yadav, J. Kishor and R.L. Yadava, “Compensation of Dielectric Cover Effects on CP Hexagonal Microstrip Antenna,” IJECET Volume 4, Issue 1, pp. 43-54, January- February (2013).

[7] R.K.Yadav, J. Kishor, and R.L. Yadava, “Design of Hexagonal Patch Antenna for Mobile Wireless System”, IEEE IJSTM, Vol. 2 Issue 4, December 2011.

[8] R.E. Collin, “Foundations for Microwave Engineering”, 2nd Edition, 484, Wiley, 2000.

[9] G. Kumar and K.P. Ray, "Broadband Microstrip Antenna", Artech House Publication, pp.1-291, 2003.

[10] T. Noro, and Y. Kazama, "A novel wideband circular polarization microstrip antenna - combination of different shaped antenna element", Antennas and Propagation Society International Symposium, 2005 IEEE Volume: 3, pp.467 – 470, 2005.

[11] Y.K. Gupta, R.K. Yadav and R.L. Yadava, "Circularly polarized truncated pentagonal shaped Microstrip patch antenna", Volume 2, International Journal of Microwaves Applications, No.6, pp.149-152, 2013.

[12] N. Zhu and W. Zhang, "A Novel Frequency-tunable Antenna with a Wide Tuning Range", Progress In Electromagnetic Research Symposium Proceedings, Stockholm, Sweden, pp. 33-38, Aug. 12-15, 2013

# SIMULATION OF MPPT TECHNIQUE USING BOOST CONVERTER FOR WIND ENERGY CONVERSION SYSTEM

**Pallavi Behera[1], D.K. Khatod[2]**

*[1]M.Tech Scholar, [2]Assistant Professor, Alternate Hydro Energy Centre,*
*Indian Institute of Technology, Roorkee (India)*

**ABSTRACT**

*The present work focuses on maximum power point tracking (MPPT) for permanent magnet synchronous generator (PMSG) based wind energy conversion systems using boost converter. Comparative results are presented for the system without MPPT and with the proposed method. Both these cases are implemented using MATLAB/Simulink®. The proposed scheme uses dc link voltage and current as inputs. The scheme has the benefits of increased system reliability and eliminates the need for shaft speed sensing. At the grid-side, an IGBT based PWM inverter is used. Simulation results are shown for a three phase resistive load under fixed and variable wind speed conditions.*

***Keywords: Boost Converter, Permanent Magnet Synchronous Generator (PMSG), Wind Energy Conversion System (WECS).***

## I. INTRODUCTION

Among the various renewable energy sources, wind has emerged as a strong contender, second only to solar energy. Wind as an energy source has great potential, particularly, in remote locations like deserts, hilly regions, and offshore areas. Since wind is a variable energy source, the use of variable speed generation scheme is an obvious choice. For small scale installations, use of PMSG is more beneficial because of its compact size, high power density, no need for gearbox [1].

To obtain maximum power from wind at any point of time, a MPPT algorithm has to be incorporated in the system. The three chief categories of tracking methods are: Tip speed ratio (TSR) control, Power signal feedback (PSF) and Hill climb search (HCS).

In TSR control, the TSR is kept constant irrespective of the wind speed by varying the generator rotor speed. The TSR is dependent on wind speed measurement which is difficult to be obtained with high accuracy and turbine shaft speed [2].

In PSF control, the data points for maximum output power and the corresponding wind turbine speed are recorded in a lookup table. The error between the optimum power and the actual turbine output power is fed to the controller for MPPT action. This method requires the knowledge of the system parameters to a high degree of accuracy [3-5].

In HCS based method, a system parameter is chosen and perturbed in small steps. This may be the rotor speed or the dc link voltage or the duty cycle of the converter, etc. Corresponding changes are noted in the output power. Depending on the sign of the derivative of power with respect to the perturbed parameter, further action

is taken to shift the operating point to a point where the slope is zero. At this operation point the power extracted is maximum [6]. This method does not require wind speed measurements and is independent of system data.

The main objective of this paper is to develop a HCS based MPPT technique using duty cycle as the perturbed parameters and analyze its application to different source conditions.

## II. MATHEMATICAL MODELING

Fig. 1 shows the block diagram of the proposed system to which MPPT method is applied. As discussed above, due to the relative advantages of PMSG, it has been utilized in the system. An uncontrolled three phase bridge rectifier has been used due to its simplicity and higher reliability. A DC/DC boost converter is used at the output terminals of the rectifier. MPPT control signal is given to the boost converter and it accordingly boosts up the voltage across the load resistor. A PWM bridge inverter is used at the load side.

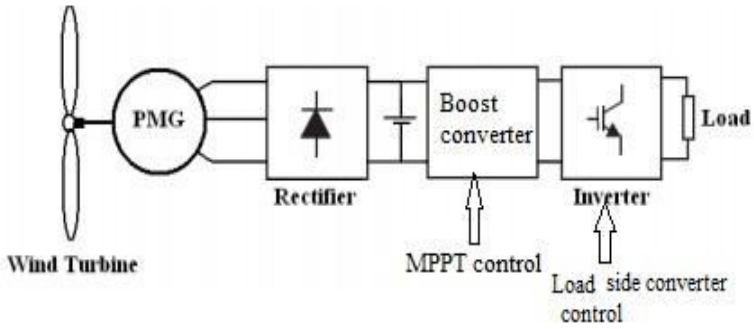



**Fig. 1 Block diagram of the WECS under study**

## III. WIND TURBINE CHARACTERISTICS

The power contained in wind $P_{wind}$ can be given as;

$$P_{wind} = \frac{1}{2}\rho A V^3 \quad (1)$$

Where, $\rho$ is the air density, $A$ is the area swept by the turbine blades, $V$ is the wind speed.

Power extracted by the wind turbine $P_m$ is;

$$P_m = \frac{1}{2}\rho A V^3 C_p(\lambda, \beta) \quad (2)$$

Where, $C_p$ is the turbine power coefficient. It is a transcendental function of blade pitch angle $(\beta)$ and tip speed ratio $(\lambda)$.

$$C_p = 0.5\left(\frac{116}{\lambda_i} - 0.4\beta - 5\right)e^{-\frac{21}{\lambda_i}} + 0.0068\lambda \quad (3)$$

$$\frac{1}{\lambda_i} = \frac{1}{\lambda} - 0.035 \quad (4)$$

$$\lambda = \frac{\omega R}{V} \quad (5)$$

Where, $\omega$ is the rotor speed, $R$ is the turbine radius.

$\beta$ is fixed by the shape of turbine blades. Thus $C_p$ can be expressed as a function of $\lambda$ only. A typical $C_p(\lambda)$ curve is shown in Fig. 2.

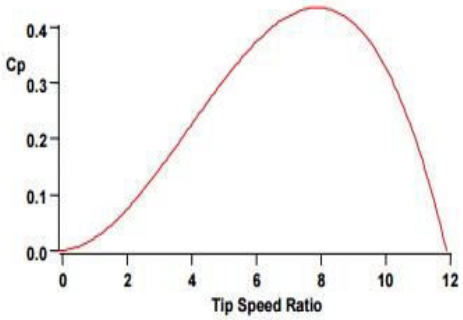



**Fig. 2 Turbine Power Coefficient as a Function of TSR [4]**

Power output of the generator $P_G$;

$$P_G = \eta P_m \quad (6)$$

Where, $\eta$ is the generator efficiency.

Using the relations (2) - (6) it can be inferred that at any particular wind speed, the generator outputs maximum power at a specific rotor speed. By controlling the duty ratio of the intermediate boost converter the load resistance referred to the generator terminals is varied along with the changes in wind, such that maximum power is extracted continuously.

The parameters of the wind turbine, generator and boost converter are given in Tables 1-3.

**Table 1. Parameters for Wind Turbine Used in Model [7]**

| Parameter | Value |
|---|---|
| Nominal mechanical output power | 8500 W |
| Base wind speed | 12 m/s |
| Maximum power at base wind speed | 7650 W |

**Table 2. Parameters for PMSG used in model [7]**

| Parameter | Value |
|---|---|
| Number of poles | 10 |
| Rated angular speed | 153 rad/s |
| Magnetic flux linkage | 0.2781 V.s |
| Stator inductance | 8.35 mH |
| Stator resistance | 0.425Ω |

## IV. BOOST CONVERTER

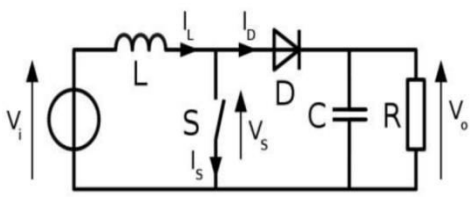



**Fig. 3 Schematic of Boost Converter [8]**

$$V_o = \frac{1}{1-D} V_I \quad (7)$$

The values for the various components of the boost converter are calculated using the following equations; [8]

$$L = \frac{V_I * (V_O - V_I)}{\Delta I_L * f_s * V_O} \quad (8)$$

$$\Delta I_L = (0.2 \; to \; 0.4) * I_O * \frac{V_O}{V_I} \quad (9)$$

$$C_{min} = \frac{I_O * D}{f_s * \Delta V_O} \quad (10)$$

Where, $V_I$ is the input voltage, $V_O$ is the output voltage, $I_O$ is the output current, $D$ is the duty ratio, $L$ is the inductance, $C_{min}$ is the minimum value of capacitance required, $\Delta I_L$ is the ripple current, $f_s$ is the switching frequency.

**Table 3. Parameters for Components of Boost Converter Used in Model**

| Parameter | Value |
|---|---|
| Inductance | 0.595 mH |
| Output capacitance | 642.8 µF |
| Switching Frequency | 40 Hz |

## V. MPPT ALGORITHM

Most of the HCS based methods use the relation between generator output power and rotor speed. These characteristics have to be stored and shaft speed measurements have to be done. The optimal output power is calculated and compared to the actual generator output power. The resulting error is used to control a power interface. Such methods require the prior information about generator characteristics, which may not be available accurately. Sensors are required for wind speed measurements which add to the cost of the entire system. As a solution to the above limitations, the proposed method is based on duty cycle of the boost converter. A detailed mathematical analysis of the used method has been presented below.

From equations (2) and (5), it can be concluded that maximum turbine output power $P_{max}$ is proportional to the cube of wind speed $V$ and hence to the cube of optimum rotor speed $\omega_{opt}$ which keeps the TSR at its optimal value $\lambda_{opt}$ for a given wind speed. Mathematically we can write,

$$P_{max} \propto V^3 \propto \omega_{opt}^3 \quad (11)$$

For a PMSG with a constant flux, the phase back electromotive force (emf) $E$ is a linear function of generator rotor speed [9], which equals the turbine speed;

$$E = K_e \varphi \omega \quad (12)$$

The phase terminal voltage $V_{ac}$ for a non-salient PMSG is written as;

$$V_{ac} = E - I_{ac}(R_s + j\omega_e L_s) = K_e \varphi \omega_{opt} - I_{ac}(R_s + j\omega_e L_s) \quad (13)$$

$$\omega_e = p\omega_{opt} \quad (14)$$

Due to the diode bridge rectifier, the ac-side voltage amplitude $V_{ac\text{-}amp}$ and the dc side voltage $V_{dc}$ can be expressed as [10];

$$V_{dc} = \frac{3\sqrt{3}}{\Pi} V_{ac-amp} \quad (15)$$

Based on equations (3.43), (3.44), (3.45) we can write the following relationship;

$$V_{dc} \propto \omega \quad (16)$$

At the point of maxima, the optimal value of the rectified dc voltage $V_{dc\text{-}opt}$ at a given wind speed is proportional to the optimal rotor speed $\omega_{opt}$.

$$V_{dc-opt} \propto \omega_{opt} \quad (17)$$

Equations (11) and (17) give;

$$P_{max} \propto V_{dc-opt}^3 \quad (18)$$

The maximum dc-side electric power at a given wind speed can be expressed as;

$$P_{dc} = \eta P_{max} = V_{dc-opt} I_{dc-opt} \quad (19)$$

$I_{dc\text{-}opt}$ is the value of dc side current at optimum point.

From (18) and (19),

$$I_{dc-opt} \propto V_{dc-opt}^2 \quad (20)$$

Or, we can write, $I_{dc-opt} = kV_{dc-opt}^2$ (21)

The relationship of equation (20) can be depicted in the form of a set of curves as shown in Fig. 4. It shows the curves of $I_{dc}$ versus $V_{dc}^2$ at different wind speeds, which are labeled as $v_{w1}$, $v_{w2}$, and $v_{w3}$ respectively. The points of intersection such as $(V_{dc1}^2, I_{dc1})$, $(V_{dc2}^2, I_{dc2})$ and $(V_{dc3}^2, I_{dc3})$ are the operating points when applying equation (21) to track MPP. The power versus $V_{dc}$ curve of these wind speeds, $vw1$, $vw2$ and $vw3$ are shown in Fig. 5. $P_1$, $P_2$ and $P_3$ are the output powers when applying MPPT.

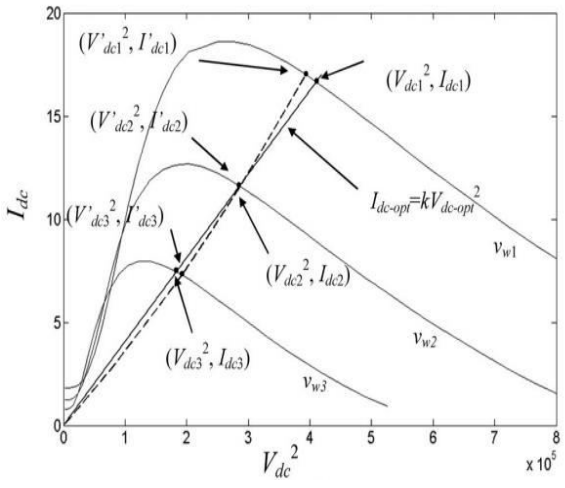



**Fig. 4 Wind Energy Electrical Characteristics – $I_{dc}$ Vs $V_{dc}^2$ Curves for Different wind Speeds [11]**

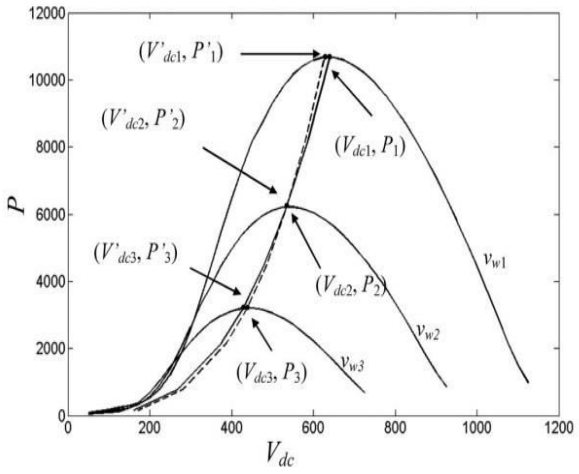



**Fig. 5 Wind Energy Electrical Characteristics – P Vs $V_{dc}$ Curves for Different Wind Speeds [11]**

Considering Fig. 5, we can observe three zones of operation:

$$\frac{dP}{dV_{dc}}_{V_{dc}=V_{dc-opt}} = 0 \tag{22}$$

$$\frac{dP}{dV_{dc}}_{V_{dc}=V_{dc-opt}} < 0 \tag{23}$$

$$\frac{dP}{dV_{dc}}_{V_{dc}=V_{dc-opt}} > 0 \tag{24}$$

Also,

$$\frac{dP}{dV} = \frac{dVI}{dV} = I + V\frac{dI}{dV} \tag{25}$$

Whenever $\frac{dP}{dV} \neq 0$ appropriate control action will be taken by the controller.

$$\text{If } \frac{dP}{dV} = 0 \text{ then } -\frac{I}{V} = \frac{dI}{dV} \tag{26}$$

For the algorithm used, the values of dc side voltages and currents are sampled at regular intervals. According to the logic described above, the duty ratio of the dc-dc converter is computed accordingly. Following this change, the rectified output voltage $V_{dc}$ of the converter changes simultaneously. The ratio I/V and ΔI/ΔV are calculated for the previous as well as current cycles and compared. The direction of search will depend on this comparison. If the change is positive and increasing, the perturbation will continue in the same direction in the subsequent cycle. This will cause the rotational speed to increase. If the change is decreasing, the direction of the search will be reversed. When the quantity $\Delta P/\Delta V$ equals zero, it means that the the MPP has been tracked and the operating point will settle around this point [12].

## VI. RESULTS AND DISCUSSION

The Simulink model of the system is shown in Fig. 6. Such a configuration can be used to supply different AC loads or connected to the grid, etc. A three phase resistive load of rating 2000 W, 415 V, 50 Hz is used for simulation. The system is simulated under fixed and variable wind speed. Output parameters are shown in blue and input parameters are shown in purple in the plots.

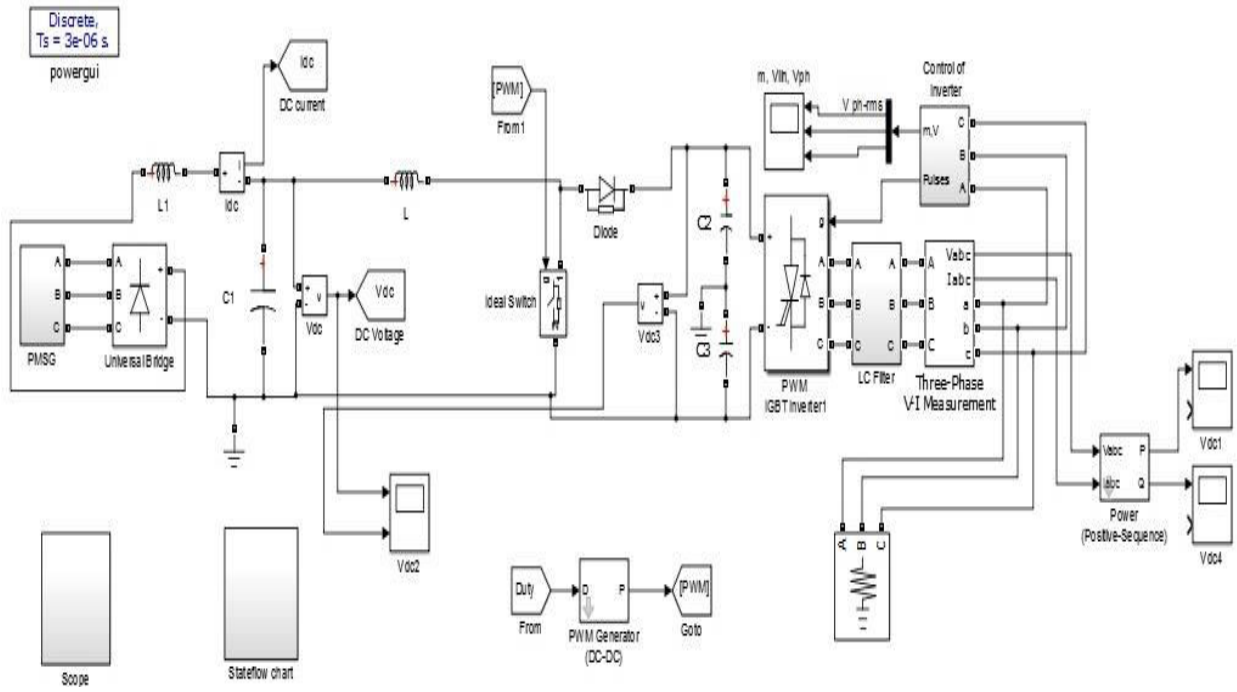



**Fig. 6 Simulink Model of the System**

## VII. FIXED WIND SPEED

Wind speed is set at 12 m/s.

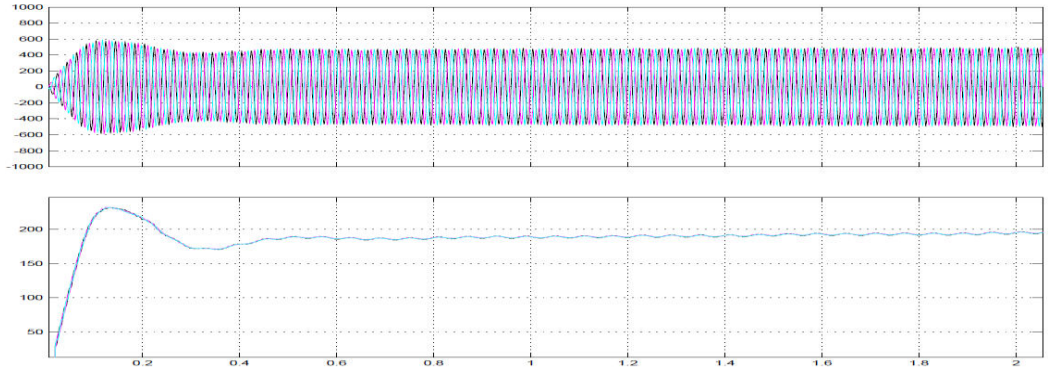


**Fig. 7 Plot of line voltage (above) and phase voltage (below) at the inverter output terminals**

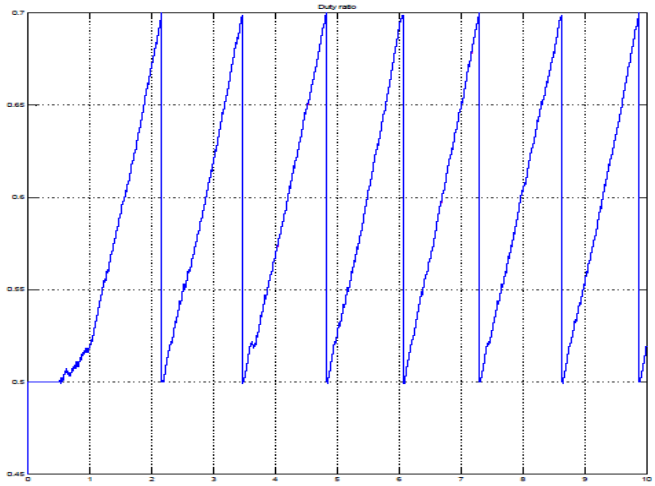


**Fig. 8 Plot of duty ratio**

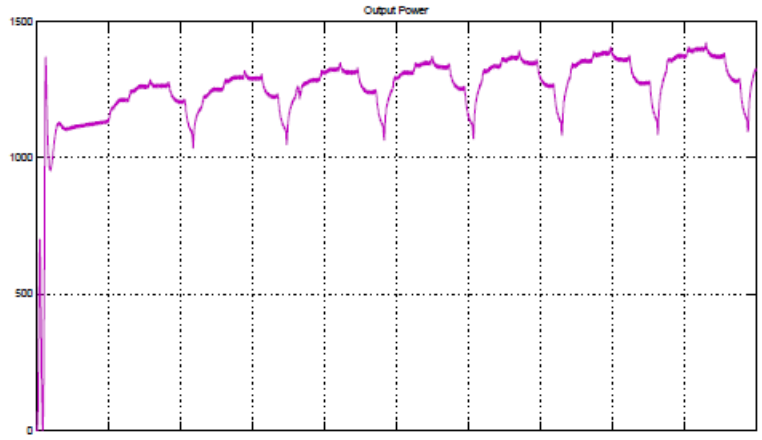



**Fig. 9 Plot of output active power**

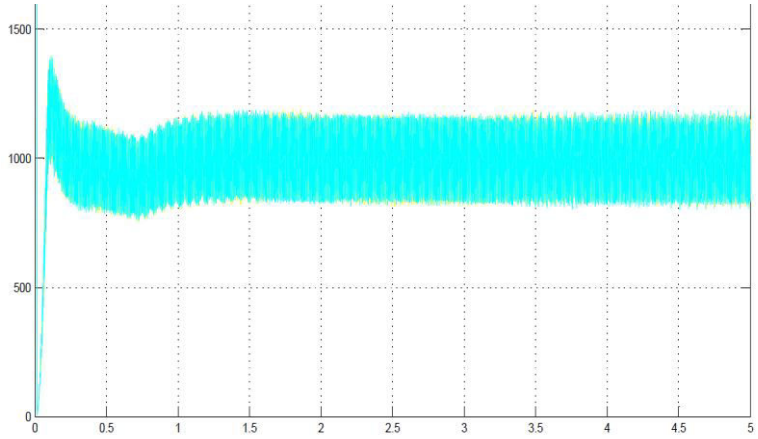



**Fig. 10 Plot of output active power without MPPT**

Fig. 7 shows the plots for output line and phase voltages at the inverter output terminals. Line voltage becomes stable around 420 V (rms). Fig. 8 shows the plot of duty ratio. Fig. 9 and 10 show the plots for output power with and without MPPT control. Without MPPT, the power delivered to load has a mean value of 1000 W. With MPPT, the output power is close to 1500 W. The results show the effectiveness of the proposed algorithm.

## VIII. STEP CHANGE IN WIND SPEED

The wind speed variation for this case is shown in Fig. 11. Fig. 12 shows the plot of output voltage. With the increase in wind speed, there is an increase in the voltage. Corresponding plots for output power with and without MPPT control are plotted in Fig. 13 and 14. Comparison for the conditions with and without power tracking control has been done. In Fig. 13, the output power is close to 1000 W when wind speed is 8 m/s and increases to nearly1300 W when speed is stepped up to 12 m/s. In Fig. 14, at a speed of 8 m/s, the output power is around 800 W and when speed is increased to 12 m/s, the power increases to around 1000 W. This shows that greater power is extracted with MPPT control.

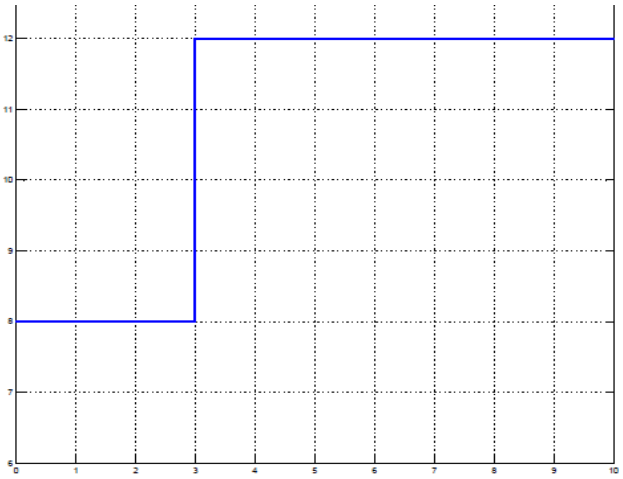



**Fig. 11 Plot of Wind Speed**

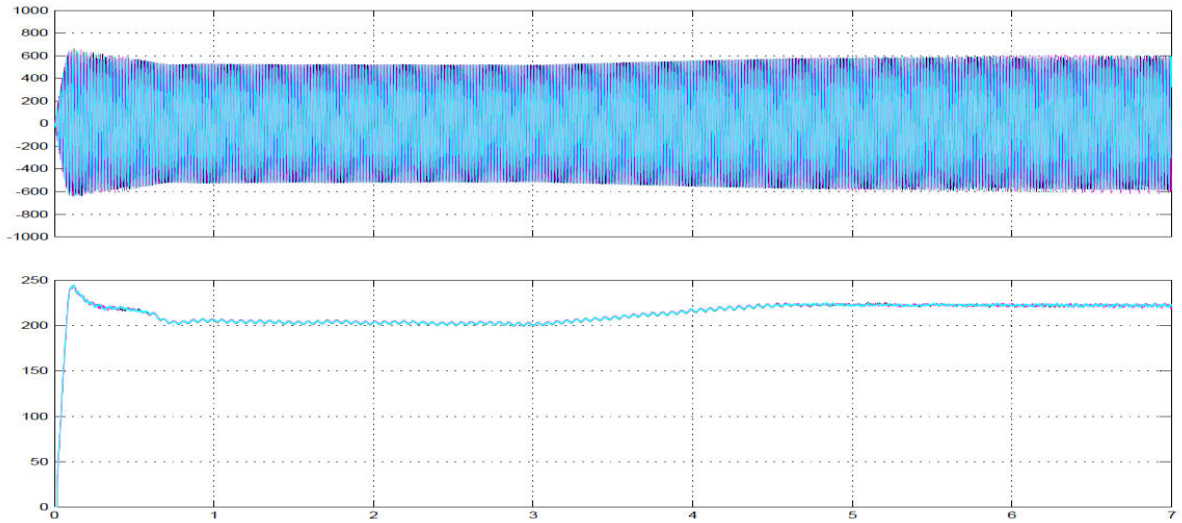



**Fig. 12 Plot of Output Line (above) and Phase (below) Voltage**

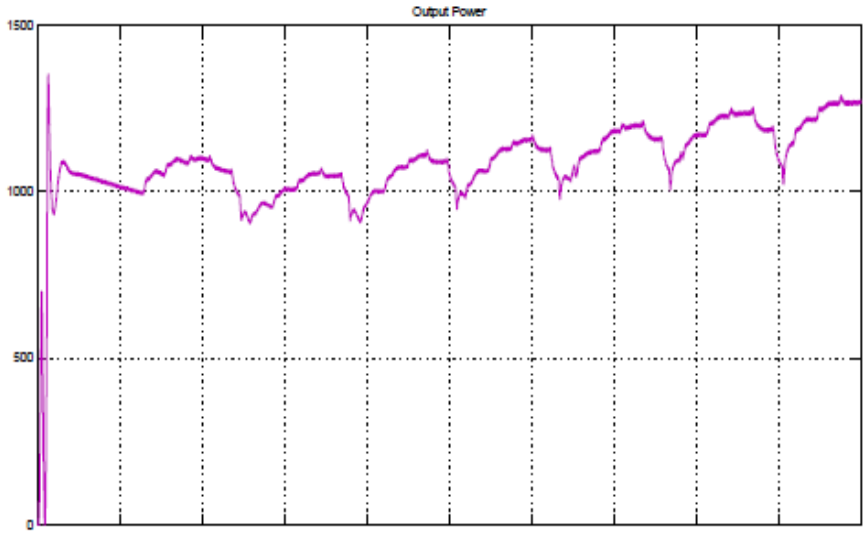



**Fig. 13 Plot of Output Active Power with MPPT Control**

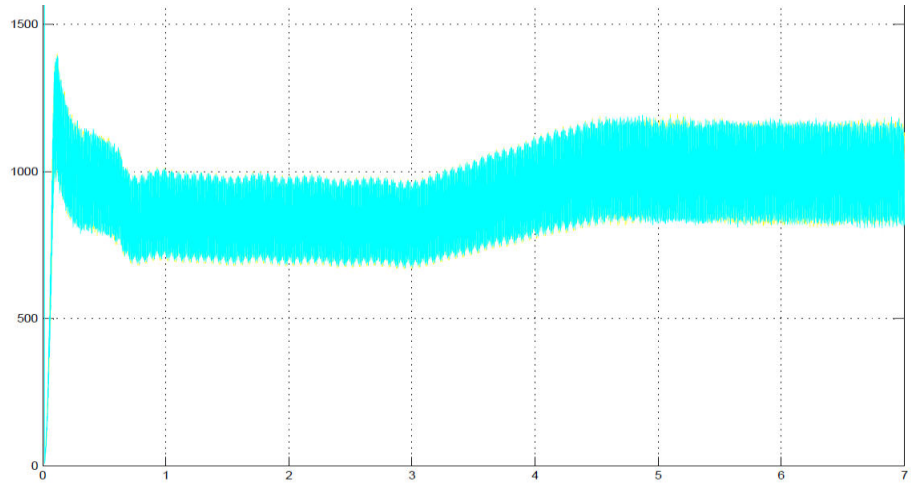



**Fig. 14 Plot of Output Active Power without MPPT Control**

## IX. CONCLUSION

A modified HCS based MPPT technique has been proposed in this paper. This technique uses dc link voltage and current as inputs. By computing the ratio of current and voltage and comparing the changes in sign, the duty ratio of the boost converter is adjusted. There is no requirement for wind speed measurements or system characteristics. Hence this method is simple in terms of implementation. The system has been simulated under fixed and variable wind speed conditions with a resistive load. Comparison has been done for the performance of the system with the proposed MPPT and without MPPT method.

The application of the discussed method can be studied for integrated renewable energy systems. Protection features can be developed to consider the circumstance of over-load or over-current.

## REFERENCES

[1] J.Puigcorbe and A.D. Beaumont, “Wind Turbine Gearbox Reliability”, 3 June, 2010. http://www.renewableenergyworld.com/rea/news/article/2010/06/wind-turbinegearbox-reliability.html

[2] Majid A. Abdullah, A.H.M. Yatim, Chee Wei, Tan, “A Study of Maximum Power Point Tracking Algorithms for Wind Energy System”, IEEE First Conference on Clean Energy and Technology CET, pp. 321 – 326, Kuala Lumpur, 27-29 June, 2011.

[3] S.M. Barakati, "Modeling and Controller Design of a Wind Energy Conversion System including Matrix Converter," PhD. dissertation, Electrical and Computer Engineering, University of Waterloo, Waterloo, 2008.

[4] K. Tan and S. Islam, “Optimum control strategies in energy conversion of PMSG wind turbine system without mechanical sensors”, IEEE Transactions on Energy Conversion, Vol. 19, Issue 2, pp. 392-399, June 2004.

[5] S. M. Barakati, M. Kazerani, and J D. Aplevich, “Maximum Power Tracking Control for a Wind Turbine System Including a Matrix Converter”, Power and Energy Society General meeting, IEEE, pp. 1, 26-30 July, 2009.

[6] B. Neammanee, S. Sirisumranukul, and S. Chatratana, “Control Performance Analysis of Feed-forward and Maximum Peak Power Tracking for Small-and Medium-Sized Fixed Pitch Wind Turbine”, 9th

International Conference on Control, Automation, Robotics and Vision (ICARCV), IEEE, pp. 1-7, Singapore, 5-8 December, 2006.

[7] A. Rolan, A. Luna, G. Vazquez, D. Aguilar and G. Azevedo, "Modeling of a Variable Speed Wind Turbine with a Permanent Magnet Synchronous Generator", IEEE International Symposium on Industrial Electronics (ISIE), pp. 734-739, Seoul, 5-8 July, 2009.

[8] B.M. Hasaneen, A.A.E. Mohammed, "Design and simulation of DC/DC boost converter", 12th International Middle East Power System Conference, MEPCON , IEEE, pp. 335-340, Aswan, 12-15 March, 2008.

[9] A. E. Fitzgerald, C. Kingsley, and S. D. Umans, "Electric Machinery", 5thedition, McGraw-Hill Publications, New York, 1992.

[10] M. H. Rashid, "Power Electronics Circuits, Devices, and Applications", Englewood Cliffs, NJ: Prentice-Hall, 1993.

[11] Y. Xia, K.H. Ahmed, and B. Williams, "A New Maximum Power Point Tracking Technique for Permanent Magnet Synchronous Generator Based Wind Energy Conversion System", IEEE transactions on Power Electronics, Vol. 26, Issue 12, pp. 3609-3620, 18 July, 2011.

[12] T. Tafticht, K. Agbossou and A. Chériti, "DC bus control of variable speed wind turbine using a buck-boost converter", Power Engineering Society General Meeting, IEEE, pp. 1-5, 2006.