International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 1, July 2014)

Design of Secured and Efficient Wireless Sensor Network with Integration to Public Cloud for Big Data Analytics

Gayathri K[1], V.Ananthanarayanan[2]

[1]*Dept of CSE, Amrita School of Engineering, Amrita University, Coimbatore.*
[2]*Assistant Professor, Dept of CSE, Amrita School of Engineering, Amrita University, Coimbatore.*

Abstract **- This paper presents a design of secured and efficient Wireless Sensor Networks with integration to public cloud for big data analytics. Now-a-days sensors are widely used in day to day life. Sensors have some limitations in terms of memory, computation, storage, communication, energy. These are the area to deal with. Cloud computing is a promising technology, which provides massive storage, computation, and software services. The limitations of wireless sensor networks are the pros of cloud computing. So by integrating these both technologies we will get greater benefits and efficiency. In this paper, we propose an integration framework of wireless sensor network with cloud, sensor data will be stored cloud and that data will be used efficiently for the needful.**

Keywords **- Cloud computing, Big Data Analytics, Wireless Sensor Networks (WSN), Services.**

I. INTRODUCTION

Wireless sensor networks is a collection of sensor nodes, which are spatially distributed to monitor environmental or physical conditions such as sound, environmental, ambient light, motion, humidity, gas, etc. Each node has a microcontroller for processing, memory, an RF transceiver for communication, a battery as a power source. Few nodes to several hundred nodes connected to a single node or several nodes to form a wireless sensor network. Sensor network provides so many uses but at the same time it poses some challenges in terms of limited storage, processing, communication and energy.

Cloud computing is anything that is hosted on the web delivered over the Internet. Cloud computing refers to applications and services offered through the Internet. These services are classified into three divisions: Infrastructure as a Service (IaaS), Software as a Service (SaaS) and Platform as a Service (PaaS). Any users with an Internet connection can share information between multiple systems or user and can access information from the cloud. The basic characteristics of the cloud are on demand self service, scalability and elasticity, resource pooling, reliability. Big data analytics is collecting complex and large amount of data that becomes difficult to process using traditional data processing applications.

It is the process of examining large amount of variety of data types to uncover the hidden patterns, unknown correlation and other unknown useful information. The key characteristics are enormous volumes of data, many types and sources of data (varied), high velocity of data, the validity of data and volatility (how long the data should be stored and valid).

II. BACKGROUND

Wireless sensor networks have limited computational power, limited memory and battery power, which leads to complexity for application developers and lot of sensed data, will get wasted. Sensors are used in so many areas such as area monitoring, environmental monitoring, military, agriculture, vehicle detection, greenhouse monitoring, pollution monitoring, etc. The sensors will detect the events according to the need and the use of sensors in that environment. For example, in greenhouse monitoring it will sense the temperature and humidity levels, if the temperature and humidity drops below a particular value it should give an alert.

Depending upon the application, sensors will sense the data and it will be reported to the base station, which can take appropriate action. Sensed data will require a different data fusion or aggregation and data propagation techniques according to the application where it is used. Sensors are also used in underwater, if there is a change in that environment, the values of sensed data will also change according to change occurred, it will detect and the change will be intimated. If the sensed data are stored and if it is analyzed, then it will be very useful to know about the natural calamities by analyzing the sensed data. But the sensors do not have more memory to save all the data which it senses and it does not have high processing power to process or analysis the sensed data.

Cloud is a model for enabling on demand network access to a shared collection of resources and services over the Internet. Large number of computers connected through a real time communication network such as Internet with the instant ability to access files and information, anywhere across the globe.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 1, July 2014)

It is to store, manage and process the data rather than in local or personal computer. It refers to delivery of computational resources over the Internet. It has three service models they are

- *Software as a Service (SaaS)* - pre-made application along with the software required, operating system, hardware and network are provided.
- Platform as a Service (PaaS) - operating system, hardware, network are provided and the user has to develop their own applications and software required for that application.
- *Infrastructure as a Service (IaaS)* - it just provides the hardware and network, the users have to install and develop its own operating system, software and applications.

The limitations of wireless sensor networks are the potential benefit of the cloud computing, such as increased storage and processing ability. By integrating both wireless sensor networks and cloud leads to greater benefits such as, high performance computing, massive storage of data, scalability, and speedy response. The main objective of this project is to expose wireless sensor networks to web services and providing the ability to seamlessly utilize cloud resources such as increased storage and processing capability.

There are some challenges in designing a Sensor-Cloud they are i) fault-tolerant ii) data transfer from senor device to server should continuous and reliable, iii) should accommodate more users to connect simultaneously, iv) users should be authenticated by different authorization rules according to their designation through web interface, v) Power (battery) is the vital issue that should be taken care because of wireless transmission would drain out the battery, vi) Service level agreement violation - cloud providers have to provide Quality of Service on user's demand, if not then it is violation of service level agreement, vii) Scalability with respect to number of subscribers.

III. Related Work

In recent years sensors are implemented in industry, agriculture, environmental protections and many other fields. In [2], the system presents an integrated wireless sensor network to monitor the information for agriculture systems namely temperature, humidity and pH values.

The purpose is to provide faster and more convenient platform for the client to obtain information to set up an agricultural system. This is suitable for any smart device which can monitor real time farmland information anywhere. The customers can access information which has Internet access to its smart device.

In this they used two techniques, distributed computing and virtualization technology. In distributed computing, they divide the whole load into smaller units. Every small work is given to a slave computer that will compute and send back the results to the master computer. There is a lack of network bandwidth, lack of storage space.

They used relational database for collection of data organized based on the relational model. To access the data from the database, we need a web service. C# is used to design the user interface and the platform used .NET that can do the integration.

In [1], proposes a system which receives and process sensor data from the wireless sensor network and supplies data services to users. It uses a cloud model which contains three layers, infrastructure layer, virtualization layer and application service layer.

Most of the healthcare system depends on their own data centers to store data, which needs high cost to maintain performance may not be reliable and provide a limited number of services. In [5], sensors are deployed in hospitals or home environment. It requires some mechanism to share the patient information available in different hospital or home environment. It uses cloud to store and allow the patient and other hospitals to access the patient information from the cloud. The information will be accessible in the community cloud, from where it can be processed by a medical professional for analysis.

In [7], proposes a novel architecture to integrate wireless sensor networks to cloud computing for improving the performance of sensor networks. In this cloud act as a virtual sink which collects data from the sensor. Data processing software is deployed in the cloud. If the size of the zone is reasonable, then the commands will reach the sensor nodes from the sink. Bandwidth is the important factor when they communicate through wireless channel. Data processing system has a master / slave architecture. Hadoop supports compression mechanism to reduce the bandwidth cost as much as possible. Master node usually connected to the Internet.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 1, July 2014)

IV. PROPOSED SYSTEM ARCHITECTURE

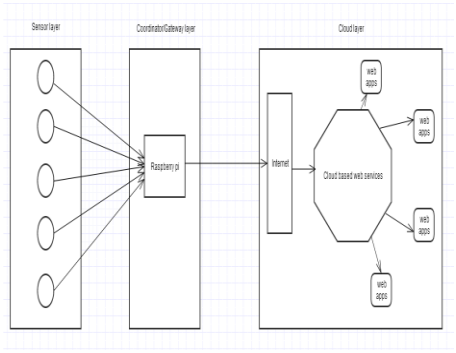

Fig1: System Architecture Outline

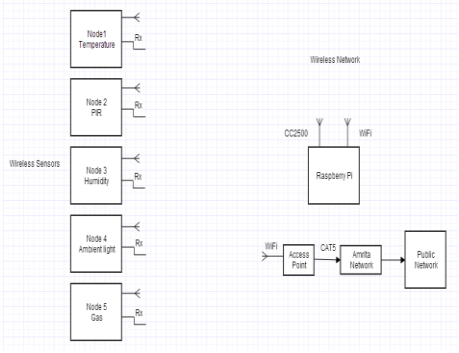

Fig 2: System Architecture

In the system architecture diagram it has been divided into three parts sensor layer, gateway layer and the cloud layer. In the sensor layer, the sensors will be deployed and it will interact with the environment to sense the data such as temperature, humidity, ambient light, etc. Sensed data will be sent to the gateway. Raspberry pi will act as the gateway it will store the sensed data, via the Internet the data will be stored in the cloud. In the cloud layer, data will be in a public cloud and also offers a web interface for end users to access the data from the cloud.

A. Node design and deployment

Arduino boards are used in this testbed. Light Dependent Resistor (LDR), temperature sensor (LM35/TMP35), Passive Infrared Sensor (PIR), humidity sensor, gas sensor (MQ135) are the sensors used. Firstly, we have to integrate these sensors onto the Arduino board. Each arduino board will be connected to all of these five sensors. After integrating these sensor nodes with the Arduino board, these boards are ready for the deployment. These should be deployed in order to achieve the performance and it should meet the requirement of the user.

The main goal may vary according to the role of the node and it is about maximizing the sensing coverage area. There are three types of coverage they are area coverage, point coverage and barrier coverage. Deploying sensors are not only based on equality of density and distribution, it's also based on other environmental factors such as climate etc.

B. Data collection of nodes

Collecting the sensor data is a critical work since sensor nodes have limited battery power and it is very critical to operate a sensor network for a long period of time. It is necessary to maximize the lime time of the nodes in the network. To collect the data sensed we are using pegasis protocol.

In Pegasis protocol, consider there are N number of nodes in the network, all the sensor nodes form a chain, according to the greedy algorithm that the sum of edges must be minimized in WSN. There are two phases:

i. Chain formation and chain head selection phase -

At this phase before each round chain head will be selected. All numbers of nodes are natural numbers between 1 to N. To choose chain head

$$CH = j \bmod N$$

If CH =0 then choose N as chain head. S(i,j) is the chain head for j^{th} round ($1 <= j <= N$).

ii. Data transmission phase

At the beginning of each round, the chain head generates 2 tokens and transmit them to 2 ends of the chain head. Two end nodes in the chain transmit in parallel. Node fuses the data and sends it to its next neighbor.

Consider there are 5 nodes S 1,1), S (1,2), S (1,3), S (1,4) and S (1,5). Consider, it is 3^{rd} round so S (1,3) is chain head. S (1,1) and S (1,5) receives token from chain head, so that they transmit their own data to S (1,2) and S (1,4) in parallel. After receiving data and token from S (1,1), node S (1,2) fuses its own data with S (1,1) data, and sends the fused data with token to S (1,3) which is the chain head. In the way it happens to nodes S (1,4) and S (1,5). Finally chain head S (1,3), fuses all data and destroys the two tokens created at the start of this phase. The following is psuedocode of Pegasis protocol:

1. Calculate chain head;
2. S (i,j) sensor node is chain head in j^{th} round;
3. S (i,j) generates two tokens and sends it to S (i,1) and S (i,N);
4. Let x=1 and y=N;

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 1, July 2014)

5. Repeat
6. If (x<j) // data transmission is on the left side of S (i, j)
7. {
8. S(i,x) fuses received data from S (i,x-1) where x>1 and its own data S (i,x);
9. S (i,x) transmits fused data and token to the neighbor node S (i,x+1);
10. x=x+1;
11. }
12. If (y>j) // data transmission is on the right side of S (i, j)
13. {
14. S(i,y) fuses received data from S (i,y+1) where y<N and its own data S (i,y);
15. S (i,y) transmits fused data and token to the neighbor node S (i,y-1);
16. y=y-1;
17. }
18. Until (x=j) and (y=j);

The main aim of pegasis protocol is to form a chain among the sensors in the network so that each sensor can send data to its near or neighbor node. Each node will transmit to and receive from its closest node. Gathered sensor data move from node to node and data aggregation will be done on the nodes which it passes through. When a node receives data it fuses that data with its data and makes it into a single message, then passes it to its closest node.

An assigned node will send all the fused data to the base station. Each and every node will get the opportunity to send data to the base station. This approach will distribute the energy load evenly among the nodes in the network. Chain building is to minimize the distance to send the data to base station, thus it reduces the energy consumption by avoiding longer distance to send data.

This protocol is able to outperform for different size of network and topologies because the overhead of dynamic cluster formation will be reduced and the number of data transmissions will be reduced by data aggregation done at each node receives data from another node. The energy load is distributed uniformly all over the network because all the nodes in the network will get a chance to act as leader, which gather all the data from the sensors and send it to the base station.

Finally the data is collected and sent it cloud via the Internet. While collecting the data itself, it is aggregated and a single message which contains all the values of all sensors which are deployed will be sent to the cloud.

C. Cloud setup and configuration

The cloud will be installed in Raspberry Pi. We installed Owncloud in Raspberry Pi. Raspberry Pi is a single board computer. Owncloud is a software system which commonly known as file hosting. It is open source and it will allow anyone to install and operate without charge on a private server, it does not have limitations on storage space or the number of clients connected.

There are some requirements for installing cloud in Raspberry Pi they are an SD card or an external hard disk, Ethernet or wireless network card and a Raspberry Pi. There are some pre-requisites for installing Owncloud in Raspberry Pi. We have to install PHP, MySQL, and Apache with SSL.

The first step is to set up the network to download and install the pre-requisites. After updation got over, start installing PHP, Apache and MySQL in Raspberry Pi.

After installation, configure PHP and restart Apache. Now we have installed and configured the pre-requisite.

Now download Owncloud, once it's downloaded it needs unzipping, then copy it in the root folder. We have to give the permission for Owncloud directory to access and setup Owncloud. In the web browser give the IP address and type Owncloud. Choose your user name and password and then click finish button below and you are done.

Login Owncloud using the username and password you have, it will say nothing is here, upload something. This page will allow us to upload files. We can upload any type of files and we can also upload music, videos, contacts, etc. the files which we are uploading to the Owncloud will be accessed by the users or clients as per their needs. We can use Java or PHP for front end user interfaces to access the data from the cloud. The cloud setup is done and the data in stored in the cloud are accessible by the user as they wish the data to be retrieved.

D. Network Setup

We have five Zigbee nodes, these nodes form a network. The coordinator is in charge of starting the network.

The following are the steps to form a Zigbee network and it is also known as a personal local area (PAN) network.

a) Radio signal search

As said before coordinator starts searching for suitable signal, that restricted to channels that are usable. While searching, it ignores the frequency on which wireless LAN is operating.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 1, July 2014)

b) PAN ID assignment

When a coordinator starts the network, it starts assigning PAN ID to the network. PAN ID can be predetermined or it can be dynamically given by knowing other network in the same frequency. The coordinator has a short address itself, by default the address is 0x0000.

c) Starting the network

The configuration of coordinator node is done by itself and starts in coordinator mode. When it starts in coordinator mode, it is prepared to respond to the queries by the other nodes that wish to join the network.

d) Acceptance and rejection of join request

The coordinator has to take decision whether a node can be permitted to join the network or not. It will have some criteria to get satisfied, if only that device is allowed to join the network. After joining that device will be allocated with an address to it.

e) Message propagation

Message propagation depends on network topology used by the network. Few network topologies are mesh topology, star topology, tree topology, etc. Message contains the final destination address if the destination node is in range. If the destination node is not in range, then the address of the nearest node which is next hop and the final destination address will be there in the message.

The intermediate nodes are not aware of the message sent and it does not know the message content.

f) Route discovery

Route discovery will be initiated when it is requested by data transmission request. A source node will send a broadcast to all nodes in the network. All the nodes eventually receive the broadcast message, one of which is our destination node. That destination node sends back a reply to the source node. When the reply travels back, it measures the hop count, signal quality. Route discovered is unidirectional.

g) Device discovery

When a node joins the network, the first work is to find out if any other node that can talk with them. The coordinator will assign addresses to all the nodes in the network and to itself, so that it can discover all the nodes in the network.

h) Operating modes of Zigbee nodes

There are three modes of operation, they are coordinator mode, router mode and end device mode. The coordinator mode will establish the network and it stores information about the network and gives security to the network. Router mode is also known as intermediate node which is used to relay data from one node to other devices. The end device mode is a low power mode, reduced functionality and cost. Supported frequency or radio frequency data rates for Zigbee nodes are 2.4GHz (250kbps), 900MHz (40kpbs) and 848MHz (20kbps).

E. Security for wireless transmission

Data should be sent securely to the receiver from the transmitter. Extended Tiny Encrypted Algorithm (ETEA) is used as base idea for security for in this project. Normally Tiny Encryption Algorithm (TEA) family is fairly strong. XTEA is more secure, has 64 block cipher and 128 keys. Keys are dynamically scheduled during runtime. XTEA uses addition for encryption and subtraction for decryption. The following code is the round function for encryption and decryption

```
for (r = 0; r < ITERATIONS; r++)
{
 y += ((z << 4) ^ (z >> 5)) + (z ^   sum) + (key[sum & 3],
z);
sum += 0x9E3779B9;
z += ((y << 4) ^ (y >> 5)) + (y ^ sum) + (key[(sum >> 11)
& 3], y);
 }

for (r = 0; r < ITERATIONS; r++)
{
y -= ((z << 4) ^ (z >> 5)) + (z ^ sum) + (key[sum & 3], z);
sum -= 0x9E3779B9;
z -= ((y << 4) ^ (y >> 5)) + (y ^ sum) + (key[(sum >> 11)
& 3], y);
 }
```

XTEA is block cipher which uses 128 bits to encrypt and decrypt the data in the block size of 68 bits. The number rounds is 64 (ITERATIONS = 64). Plain text is given in round 1 after 64 rounds it is converted into decrypted text by using the logic described above.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 1, July 2014)

In wireless networks the data will be transmitted in the form of frames. The frame contains start delimiter, length of the data, destination address, optional bytes, original data and checksum. For wireless transmissions, I considered the attacks in three perspective, they are i) Unknown device ID, ii) Unknown frequency and iii) Unknown key.

In case (i) we connect upto 256 devices in a network but there are so many modes like idle, active, park etc, so in a network approximately we can have 200 devices connected. Each and every device will be having its unique id. Sender device will have the receiver address to which device it has to send the data, this address will mentioned inside the frame. We encrypting the whole data i.e., frame, so device ID cannot found easily by the intruder, it has the probability 200 trails to found out the device ID.

In case (ii) we have 1 to 200 channels for data transmission, the intruder have to scan the channel from 1 to 200 to know in which channel the data is transmitting to get the data. So it also has the probability of 200 trails to know the frequency. In case (iii) we have 128 bits of key, if the intruder wants to find the key, he/she has to check for 2^{128}.

Our algorithm will do the following steps, first the frame is taken as plain text. It is divided into four parts and each part is encrypted separately and all the four parts of data will be encrypted. These four parts of the frame is sent separately in different channel, in such a way that if the intruder can snoop and found the channel on which data is sent, he/she cannot get the whole data. Because data is divided into four parts, so intruder can get only a part of data, which is of no use. The transmitter and receiver will be in same frequency to receive the data to decrypt it. After decrypting it will check for the checksum, if the checksum is same then the received data is reliable.

V. Result Analysis

A. Performance evaluation

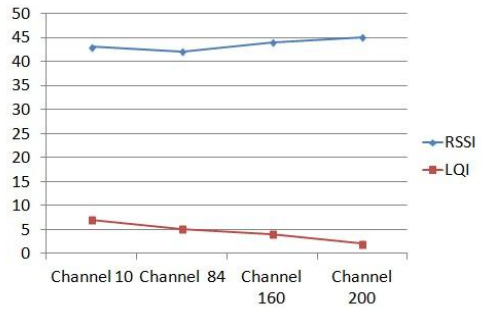

Figure 4 : RSSI vs LQI

From the above mentioned graph we can know that when the received signal strength increases the link quality will also get increased.

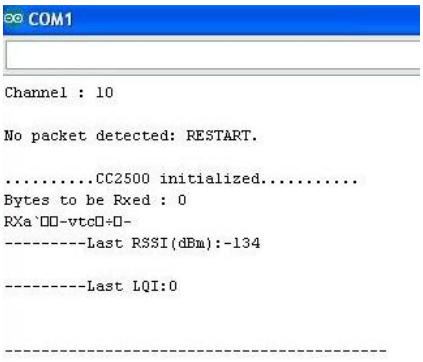

Figure 5 : Intruder screen

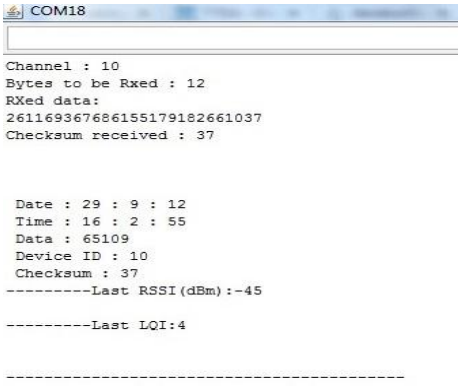

Figure 6 : Receiver screen

When the intruder is trying to access data from the transmitter, it cannot break the frame which is encrypted. Since device ID of the receiver is not matching with the intruder device ID and the frequency of transmitter and receiver does not match with intruder's frequency, also decryption key in receiver and intruder is not equal. Hence garbage value is displayed in Fig 5 intruder screen and communication cannot be made.

Since receiver matches all the cases i.e., device ID, frequency and key, it can be able to decrypt and original information is received successfully.

VI. Conclusion And Future Work

Wireless sensor networks are extensively used in data gathering places such as vehicle monitoring, pollution monitoring, health monitoring, with fast development of its growth it faces so many problems such as how to efficiently use the sensor data and fully utilize them.

In this paper, we presented the design of secured and efficient WSN with integration to public cloud for big data analytics. In this framework, sensor nodes are connected to form a network where Zigbee is used to route the packets from sensor nodes to the middleware and the temperature values are stored in the system as a text file. That file will be uploaded to the cloud server by synchronizing cloud server to the client. By the help of web application, the customer can send queries to the server and the server will send the information send by the server.

Acknoledgement

I am extremely grateful to Mr. V. Ananthanarayanan, Assistant Professor, Department CSE., Amrita School of Engineering for his sincere guidance, inspiration and right direction throughout the project. I express our wholehearted indebtedness to faculty and staffs of Department of Computer Science Engineering, Amrita School of Engineering, Coimbatore for providing the necessary system requirements and moral support.

REFERENCES

[1] Pengfei You, Huiba Li, Yuxing Peng and Ziyang Li, An Integration Framework of Cloud Computing with Wireless Sensor Networks, Lecture Notes in Electrical Engineering, Volume 214, Springer 2012.

[2] Wen-Yaw Chung, Pei-Shan Yu, Chao-Jen Huang, Cloud Computing System Based on Wireless Sensor Network, IEEE Proceedings of the 2013 Federated Conference on Computer Science and Information Systems pp. 877–880.

[3] Heikki Mahkonen, Teemu Rinta-aho, Tero Kauppinen, Mohit Sethi, Jimmy Kjällman, Patrik Salmela, Tony Jokikyyny, Demo: Secure M2M Cloud Testbe, MobiCom'13, September 30–October 4, 2013, Miami, FL, USA, ACM 978-1-4503-1999-7/13/09.

[4] Ahmed K, Gregory M, Integrating Wireless Sensor Networks with Cloud Computing Mobile Ad-hoc and Sensor Networks (MSN), IEEE Seventh International Conference on 2011, Page(s):364 – 366, ISBN:978-1-4577-2178-6.

[5] Perumal.B , Pallikonda Rajasekaran.M and Ramalingam.H.M, WSN Integrated Cloud for Automated Telemedicine (ATM) Based e-Healthcare Application, 2012 4th International Conference on Bioinformatics and Biomedical Technology, IPCBEE vol.29 (2012) © (2012) IACSIT Press, Singapore.

[6] Rajeev Piyare and Seong Ro Lee, Towards Internet of Things (IOTS): Integration of Wireless Sensor Network to Cloud Services for Data Collection and Sharing, International Journal of Computer Networks & Communications (IJCNC) Vol.5, No.5, September 2013.

[7] Peng Zhang, Zheng Yan, Hanlin Sun, A Novel Architecture Based on Cloud Computing for Wireless Sensor Network, Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013).

[8] Zhixiang Yuan, Jinxiang Cheng, The Design and Realization of Wireless Sensor Network Gateway Node, Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013).

[9] Atif Alamri, Wasai Shadab Ansari, Mohammad Mehedi Hassan, M. Shamim Hossain, Abdulhameed Alelaiwi, and M. Anwar Hossain, A Survey on Sensor-Cloud: Architecture, Applications, and Approaches, Hindawi Publishing Corporation International Journal of Distributed Sensor Networks, Volume 2013, Article ID 917923, 18 pages.

[10] Sanjit Kumar Dash, Subasish Mohapatra and Prasant Kumar Pattnaik, A Survey on Applications of Wireless Sensor Network Using Cloud Computing, International Journal of Computer Science & Emerging Technologies (E-ISSN: 2044-6004), Volume 1, Issue 4, December 2010.

[11] Introduction to Cloud Computing, Copyright © 2010 Dialogic corporation, 07/10 12023-01.

[12] Ouadoudi Zytoune and Driss Aboutajdine, A Lifetime Extension Protocol for Data Gathering in Wireless Sensor Networks, ISSR Journal 2013.

[13] Poulami Dutta, A Secure Hierarchical Protocol for Wireless Sensor Networks, AJER e-ISSN : 2320-0847 p-ISSN : 2320-0936 Volume-03, Issue-02, pp-191-203, 2014.

[14] Rihac Mehta, O.S. Khanna, Reducing Chain Complexity using Honey Bee Optimization in Wireless sensor network, International Journal of Computer Trends and Technology (IJCTT) - volume4Issue4 –April 2013.

[15] Zheng Gengsheng, Hu Zhengbing, A Clustering Protocol Using Multiple Chain Strategy in WSNs, Journal of Networks, Vol.5, No.5, May 2010.