

# **A Framework for Efficient Query Processing in Wireless Sensor Networks**

A Thesis submitted in fulfillment of the requirement for the award of the degree of

**DOCTOR OF PHILOSOPHY**

In

**Computer Science and Engineering**

*Submitted by*

**Vandana Jindal**

(Registration No. 950803012)

Under the Supervision of

**Dr. A. K. Verma**

**Associate Professor**

**Dept. of Computer Science and Engineering  
Thapar University, Patiala (Punjab)**

**Dr. Seema Bawa**

**Professor**

**Dept. of Computer Science and Engineering  
Thapar University, Patiala (Punjab)**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

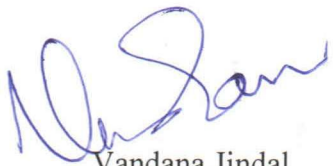
**THAPAR UNIVERSITY, PATIALA-147004 (INDIA)**

**JULY 2016**

## Certificate

I hereby certify that the work which is being presented in the thesis entitled "A Framework for Efficient Query Processing in Wireless Sensor Networks" in partial fulfillments of the requirements for the award of the degree of doctor of philosophy and submitted in the Department of Computer Science and Engineering of Thapar University, Patiala is an authentic record of my own work carried out during a period from July 2009 to July 2016 under the supervision of Dr. A.K. Verma, Associate Professor, CSED,Thapar University, Patiala and Dr. Seema Bawa, Professor, CSED,Thapar University, Patiala.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute/ University.




Vandana Jindal  
(Regn No. 950803012)

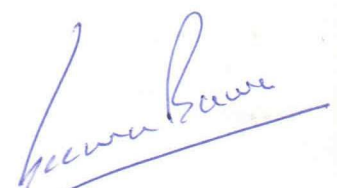
---

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

Date: 4<sup>th</sup> July 2016



Dr. A.K. Verma



Dr. Seema Bawa

## Abstract

Though Wireless Sensor Networks (WSN) have invaded every walk of modern life due to low setup time and cost yet Research and Development (R&D) for further reduction in sensor costs and energy consumption is still a high priority area for research community. Major chunk of energy consumption of WSN is in communication activity therefore most of the research has targeted the aspect of reduction in energy consumption during communication of data. As a result of research efficient hardware and numerous network protocol stacks along with data reduction techniques have been developed.

In a WSN, reduction of data is possible at two stages: 'In-network' and at 'Base Station' (BS). Due to limited processing capacity of sensor nodes, data processing in network is possible to a limited extent only whereas, at the BS the room for manipulation is much more.

In this work processing at both the stages In-network aggregation and BS optimization has been used to bring about reduction in data to be communicated.

The proposed framework: "Compression At input with Multi query Optimization at Sink" (CATMOS) is built upon query type data extraction technique and makes use of both In-network optimization and Base Station optimization for efficient Query Processing. In the work, element of Network management has been employed by using In-network compression. Query optimization at BS, has been achieved by merging new queries with already running queries resulting into fewer new queries. Query language syntax has also been modified in which single syllable Static variables have been used in place of repeated phrases. The framework has been designed to benefit any WSN irrespective of application being served by it. The work has adopted a simple generalized approach for efficient operation of WSN, therefore, it is different, from most of the researches which has used highly technical or specialized means and have intended to be application specific.

This proposed generalized approach is applicable to WSN of any size or topology. It isolates the user from the lower level details of network configuration or protocols used.

In 'In-network aggregation' the data has been compressed at the node level. The main objective has been to reduce the data size by eliminating spatio-temporal redundancies. Data

fidelity has been given importance therefore quantitative reduction in data has been allowed without any compromise in quality of data. Reduced data to be communicated has produced desired result of less consumption of sensor node energy. Three compression algorithms i.e., Huffman, LZW and Deflate compression algorithms have been examined on a Sun Small Programmable Object Technology (SunSPOT) network. Possible data reduction through each algorithm has been tested on a simulator ‘Solarium’.

For BS optimization, a frugal approach has been adopted. In case of any new query, data available in response to already running queries within the network has been scrutinized to check whether the reply to the newly injected queries can be generated from the data or not. If at all, query has to be sent to the network, it has been merged with the running queries. A novel algorithm for rewriting queries after merging two or more queries has been proposed. New queries are rewritten along with the already running queries within the network. Only those mergers have been chosen for action, which provide actual Gains as verified through the algorithm. The merged queries have been given the name ‘Synthetic Queries’. The proposed algorithm sees to it that the chosen synthetic queries achieve all the desired objectives. In addition to this query language syntax modification by using single syllable static variable has also been used to achieve additional Gains.

Simulator results of compression and query optimization have been encouraging. More than 10% Gain has been possible through use of these simple compression algorithms only.

**Keywords:** Wireless Sensor Network (WSN), Compression at Input with Multi-query Optimization at Sink (CATMOS), Multiple Query Optimizations (MQO), In-network processing, Base Station optimization, SunSPOT

## **Acknowledgements**

I would like to express my gratitude to my guides Dr. Anil Kumar Verma and Dr. Seema Bawa. They have guided and supervised me in my pursuit of knowledge which has come up in the form of the present work. I thank them for encouraging me at every step which helped me to complete this task. Their advice on both research as well as on my career has been priceless. My deep regards to Dr. P. Gopalan, Director, TU for all the facilities provided to me during my research work. I would also like to thank my committee members Dr. Maninder Singh and Dr. Rajesh Khanna and Dr. O. P. Pandey, Dean of Research & Sponsored Projects, for their meaningful comments and useful suggestions during my research which helped me a lot in improving my shortcomings.

I will also like to thank my parents for their support. Their blessings and prayers for me gave me strength throughout this long journey.

I would also like to thank all of my friends especially Mr. Sandeep Verma, Prof. Inderjit Singh and Prof. Shallu who supported me in writing, and encouraged me to strive hard towards my goal.

Mere words cannot express my feeling of debt and gratitude to my daughter Aayushi Raj and son Aaryan Raj for allowing me the time which I had to devote to my work meant for them. At the end I would like to express appreciation to my husband R.K. Jindal, who has always been my support.

Vandana Jindal

# Table of Contents

Candidate's Declaration	i
Abstract	ii
Acknowledgements	iv
Table of Contents	v
List of Abbreviations	vi
List of Tables	viii
List of Figures	ix
Thesis contributions	xi
List of Publications	xii
<b>Chapter 1: Introduction</b>	<b>1</b>
1.1 History of Wireless Sensor Networks (WSN)	2
1.2 Features of WSN	5
1.3 WSN	5
1.3.1 Sensor Node Architecture	8
1.3.1.1 Network Layer Stack of WSN	9
1.3.1.2 Network Components	11
1.3.2 WSN Infrastructure	15
1.3.3 Attributes of WSN	16
1.3.4 Applications of WSN	17
1.4 Types of WSN	21
1.5 The Underlying Technology	23
1.6 Query Processing	25
1.6.1 Query Types	26
1.6.2 General Strategy for Query Processing	28
1.6.3 Relationship between Query processing and WSN	28
1.6.4 Query Processing in WSN	29
1.6.5 Multiple Query Optimizations (MQO) in WSN	29
1.6.5.1 Key requirements of MQO	31
1.7 Metrics for WSN Evaluation	32
1.8 Gaps	34
1.9 Problem formulation	35
1.10 Research Objectives	35
1.11 Thesis Organization	36

<b>Chapter 2: Literature Review</b>	<b>38</b>
2.1 Agent based approach	41
2.2 Macro programming based approach	42
2.3 Query based approach	42
2.4 DBMS Query Processing vs. WSN Query Processing	44
2.5 Energy Efficient Approaches for WSN	45
2.5.1 Mobility based	46
2.5.2 Duty Cycling	47
2.5.3 Database Optimization	48
2.5.3.1 Data Acquisition	48
2.5.3.2 Data Processing	51
2.6 Types of Query processing Frameworks	57
2.7 Summary	64
<b>Chapter 3: Proposed Framework for Query processing in WSN: Compression at Input with Multi-query Optimization at Sink ‘CATMOS’</b>	<b>65</b>
3.1 Components of the Proposed Framework	66
3.1.1 WSN Structure	66
3.1.2 Base station/ Sink	67
3.1.3 Sensor nodes/ Motes	68
3.1.4 User	68
3.1.5 Software	68
3.1.6 SPOT nodes and Software Tools	68
3.2 Query processor architecture	70
3.3 Working of CATMOS	71
3.3.1 Proposed Strategy for Energy Optimization	74
3.4 Summary	86
<b>Chapter 4: Simulation Setup, Results and Analysis</b>	<b>87</b>
4.1 Simulation Setup	87
4.2 Test cases, test results and test analysis	93
4.3 Summary	101
<b>Chapter 5: Conclusion and Future Scope</b>	<b>102</b>
<b>References</b>	<b>105</b>

# List of Important Abbreviations

## Abbreviations/

### Acronyms

### Description

ACQUIRE	Active Query Forwarding
ADC	Analog-to-Digital Convertors
ARPANET	Advanced Research Project Agency
ANN	Artificial Neural Network
ACO	Ant Colony Optimization
BECAN	Bandwidth Efficient Cooperative Authentication
BS	Base-Station
CATMOS	Compression at Input with Multi-query Optimization at Sink
DARPA	Defense Advanced Research Projects Agency
DCQS	Dynamic Conflict-free Query Scheduling
DML	Data Manipulation Language
DQT	Distributed Quad-Tree
DSN	Distributed Sensor Networks
DSWare	Data Service middleWare
ED	Event Detection
FH	Fixed Host

## List of Important Abbreviations Contd.

Fjord	Framework in Java for Operation on Remote data Streams
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IoT	Internet of Things
ISA	International Society for Automation
ISM	Industry Scientific and Medical
LEACH	Low Energy Adaptive Clustering Hierarchy
LZW	Lempel–Ziv–Welch
MAC	Medium Access Control
MEMS	Micro Electro Mechanical System
MH	Mobile Host
MQO	Multiple Query Optimization
NOAA	National Oceanographic and Atmospheric Administration
OS	Operating System
OTAP	Over The Air Programming
QoS	Quality-of-Service
SDK	Software Development Kit
SPE	Spatial Process Estimation
SOSUS	Sound Surveillance System

## List of Important Abbreviations Contd.

SQTL	Sensor Query and Tasking Language
SunSPOT	Sun Small Programmable Object Technology
TAG	Tiny Aggregation
TTMQO	Two-Tier Multiple Query Optimization
VM	Virtual Machine
VSN	Virtual Sensor Network
WPAN	Wireless Personal Area Network
WSN	Wireless Sensor Network

## List of Tables

Table No.	Description	
2.2	DBMS Query Processing vs. WSN Query Processing	44
3.5	Table of Notations	76
4.6	Raw data (bytes) and Compressed Data (bytes) using Huffman/ LZW/ Deflate algorithms	86
4.7	Compression Factor (CF) (%age compression using Huffman/ LZW/Deflate algorithm)	87
4.8	Percentage compressions (Huffman / LZW Compression algorithm)	87
4.15	Comparative features of various frameworks with CATMOS	101

# List of Figures

<b>Figure No.</b>	<b>Description</b>	
1.1	Technology transitions in WSN	3
1.2	WSN gain market traction with decrease in sensor costs	4
1.3	Abridgement of WSN	6
1.4	A Wireless Sensor Network comprising of Sensor nodes and Base Station	7
1.5	Comparison between the OSI Model and the TCP Model	9
1.6	Network layer Stack of WSN	10
1.7	Sensor node Architecture & tasks performed by various Components	13
1.8	Anatomy of a SunSpot	14
1.9	A WSN comprising of Base-Station and free Motes	15
1.10	Applications of WSN	18
1.11	Underground WSN	21
1.12	Underwater WSN	22
1.13	Multimedia WSN	23
1.14	Wireless sensor communication technologies comparing data rate and range	24
1.15	Various query types	27
1.16	Categories of Query optimization methods (Base station optimization and In-Network optimization)	30
2.1	Various Information extraction techniques	39
2.3	Various Energy Efficient approaches in a WSN	45
2.4	Various Database Optimization approaches in a WSN	47
2.5	Various Adaptive Sensing/ Data Acquisition approaches in a	

## List of Figures Contd.

	WSN	49
2.6	Various Energy Efficient Data Processing approaches in a WSN	50
2.7	Data Fusion- Collecting results from multiple nodes	53
2.8	Data aggregation performed by Sensor Node Coordinator	54
2.9	Various Energy Efficient Database Optimization approaches in a WSN ata glance	55
2.10	TinyDb approach: Query Processing Model	57
2.11	The COUGAR approach	58
2.12	Fjord’s architecture	60
3.1	A multi hop structure of a WSN	66
3.2	Sequence diagram of Query Request	70
3.3	Sequence Diagram of sensed data collection	71
3.4	The proposed architecture “CATMOS”	75
3.6	In-network processing	79
3.7	Base station optimization	83
4.1	Sun SPOTs with an AA battery	87
4.2	The SPOT Manager	88
4.3	NetBeans	89

## List of Figures Contd.

4.4	Solarium Application	89
4.5	Solarium showing the graph from the sensed data	90
4.9	Raw Data vs. Compressed Data using Huffman Compression algorithm	92
4.10	Raw Data vs. Compressed Data using LZW algorithm	94
4.11	Raw Data vs. Compressed Data using Deflate algorithm	95
4.12	Results obtained from Number of Sensor readings transmitted and received	96
4.13	Compression Factor (without Static Variable) vs. Compression Factor (with Static Variable) using Huffman Compression algorithm	98
4.14	Compression Factor (without Static Variable) vs. Compression Factor (with Static Variable) using LZW algorithm	99
4.15	Comparative features of various frameworks with CATMOS	101

## **Thesis contributions**

Literature covering various aspects of Query Processing in WSN was studied and it was observed and analyzed that query is the most general and easy approach to extract information from a WSN. It is also established from the literature that optimization of communication load is the most effective means for achieving energy efficiency. Therefore, query optimization was chosen as the area of research work.

Based upon the literature review, elimination of redundant data may be done with the help of two stages in In-network and at Base Station (BS). Due to limited processing capacity of sensor nodes, the In-network data processing is possible to a limited extent only whereas, at the BS the room for manipulation is much more. In this stage, the response to any query for the available data at BS is scrutinized to find out whether the query can be replied or not from the data.

For query optimization multiple queries merging at base station through an iterative algorithm was proposed. Various queries are merged and iterative algorithm is employed to select the best merger based on the maximum 'Gain' achieved. In-network processing can also help in eliminating redundancy in sensor data and thus, reducing the communication load. Therefore, compression algorithms suitable for WSN sensor nodes were selected and used on data streams of WSN.

The proposed framework 'CATMOS' employs In-network processing, multiple query optimizations at base station and modification of query syntax. 'Solarium' the most general simulator for SunSPOT motes was used for experimental verification of the proposed framework CATMOS.

## List of Publications

### Papers published in Refereed Journals:

- [1] Vandana Jindal, Anil Kumar Verma, Seema Bawa, “**Survey on Query Processing & Optimization Techniques in WSN**”, *International Journal of Computer Science and Information Security, (IJCSIS)* , Vol. 14, No. 2, February 2016, pp.54-59
- [2] Vandana Jindal, Anil Kumar Verma, Seema Bawa, “**Quantitative Reduction in Communication Load for Energy Efficiency in WSN**”, *Wireless Personal Communications (WPC)*, Vol. 83, Number (4), August (II) 2015, pp. 1-16. (SCI-E) (Impact Factor: 0.693)
- [3] Vandana Jindal, Anil Kumar Verma, Seema Bawa, “**Green WSN- Optimization of Energy Use through Reduction in Communication Workload**” *International Journal in Foundations of Computer Science & Technology (IJFCST)*, Vol.5, No.3, May 2015, pp. 57-69.
- [4] Vandana Jindal, Anil Kumar Verma, Seema Bawa, “**Current Scenario in WSNs**”, *International Journal of Advanced Research in Computer Engineering and Technology (IJARCET)*, vol.1(4) Jun 2012, pp 291-294.
- [5] Vandana Jindal, Anil Kumar Verma, Seema Bawa, “**Query Processing**”, *International Journal of Computer Theory and Engineering (IJCTE)(Academy Publishers)*, vol. 3(3), 2011, pp.440-443.

### Papers published in International Conference Proceedings:

- [1] Vandana Jindal, Anil Kumar Verma, Seema Bawa, “**Two pronged Strategy for Energy Optimization in WSNs by using In-network Compression and Synthesis of Multiple Queries at Base-Station**”, *In Proceedings of the International Conference on Systems, Control, Signal Processing and Informatics (SCSI 2015), Barcelona, Spain, April 7<sup>th</sup> -9<sup>th</sup> , 2015, pp 135-141.*
- [2] Vandana Jindal, Anil Kumar Verma, Seema Bawa, “**Applications and Challenges of Ubiquitous Sensing**”, *International Conference on Information and Mathematical Sciences (IMS-13), Bathinda, Oct 24<sup>th</sup> -26<sup>th</sup> , 2013(Procedia Elsevier)*

# **Chapter-1**

## **Introduction**

Wireless Sensor Network (WSN) is a distributed network of a large number of randomly deployed tiny devices called nodes which have inbuilt computational, storage and communication capabilities. WSN is supposed to operate in an unattended mode and collect data of interest from its coverage area. The nodes are low cost devices also prone to failures due to harsh and uncertain environment in which these operate. Though the nodes are self configuring and network reconfigures itself on failure of nodes however high failure rate may affect the network badly. In addition to failure due to environmental conditions, exhaustion of limited battery also renders a node useless. As the life of a WSN depends upon life of its constituent nodes, the main objective in a sensor network is to extend life of a node. Environmental conditions cannot be changed as these are to be monitored by the WSN, so node failure due to battery depletion is to be minimized. Out of the three basic operations-Sensing, Computation and Communication carried out by a sensor node, communication is the most energy consuming operation. Therefore, to extend life of the battery of a sensor main thrust should be on reduction of communication load. Endeavour to use limited battery power of a node efficiently has led to many hardware and software advances in WSN area.

Adoption of new fabrication technologies has improved operating efficiencies of these nodes. Besides the research in hardware improvements, research community is more

inclined towards research at software level as it has many possibilities and great potential to yield significant results. Data collected by sensors is extracted by various applications which set up data flow or communication of data in the network. Various strategies such as application specific MAC and routing protocols, In-network processing, data aggregation techniques etc. have been employed to minimize communication. Query is the most popular and convenient method of data extractions. Optimization of the queries at the Base Station (BS) is another tool to reduce communication load. Modern WSN are so versatile that they can serve simultaneously multiple users having varying informational interests.

A framework has been proposed under which similar queries fired by various users at the BS are merged and synthesized into new queries. An algorithm has been developed to check whether the merger is beneficial or not, for which a new metric 'Gain' has been introduced. Besides this In-network compression has also been employed, as a result of which there is a considerable reduction in communication load. Compression results have been compared by employing some well known compression algorithms.

This chapter gives an overview of WSN, building blocks, its architecture, characteristics and applications and information extraction from WSN through various means. Queries – the most popular means of information extraction and their processing has also been discussed. Various performance evaluation metrics for WSN and sensor nodes have also been defined.

## **1.1 History of Wireless Sensor Networks (WSN)**

Though rapid interest and research in WSN fields have taken place only recently but, use of sensors for specialized services is not new. During the Cold War, quiet Soviet submarines were detected by deploying the Sound Surveillance System (SOSUS), which employed acoustic sensors. These systems are now adopted by National Oceanographic and Atmospheric Administration (NOAA) for sensing the events in the oceans. Simultaneously, Air defense radar networks were developed employing aerostats as

sensors. The predecessor to the internet, Advanced Research Project Agency (ARPANET) formed by US DARPA in 1969, served as a test bed for new networking technologies connecting various universities and research centers. A sensor network can be assumed to have many spatially distributed autonomous sensing devices which route the information to a node which can make the best use of the acquired information. The actual WSN may be traced back to the Distributed Sensor Networks (DSN) program which started in 1980 at Defense Advanced Research Projects Agency (DARPA).

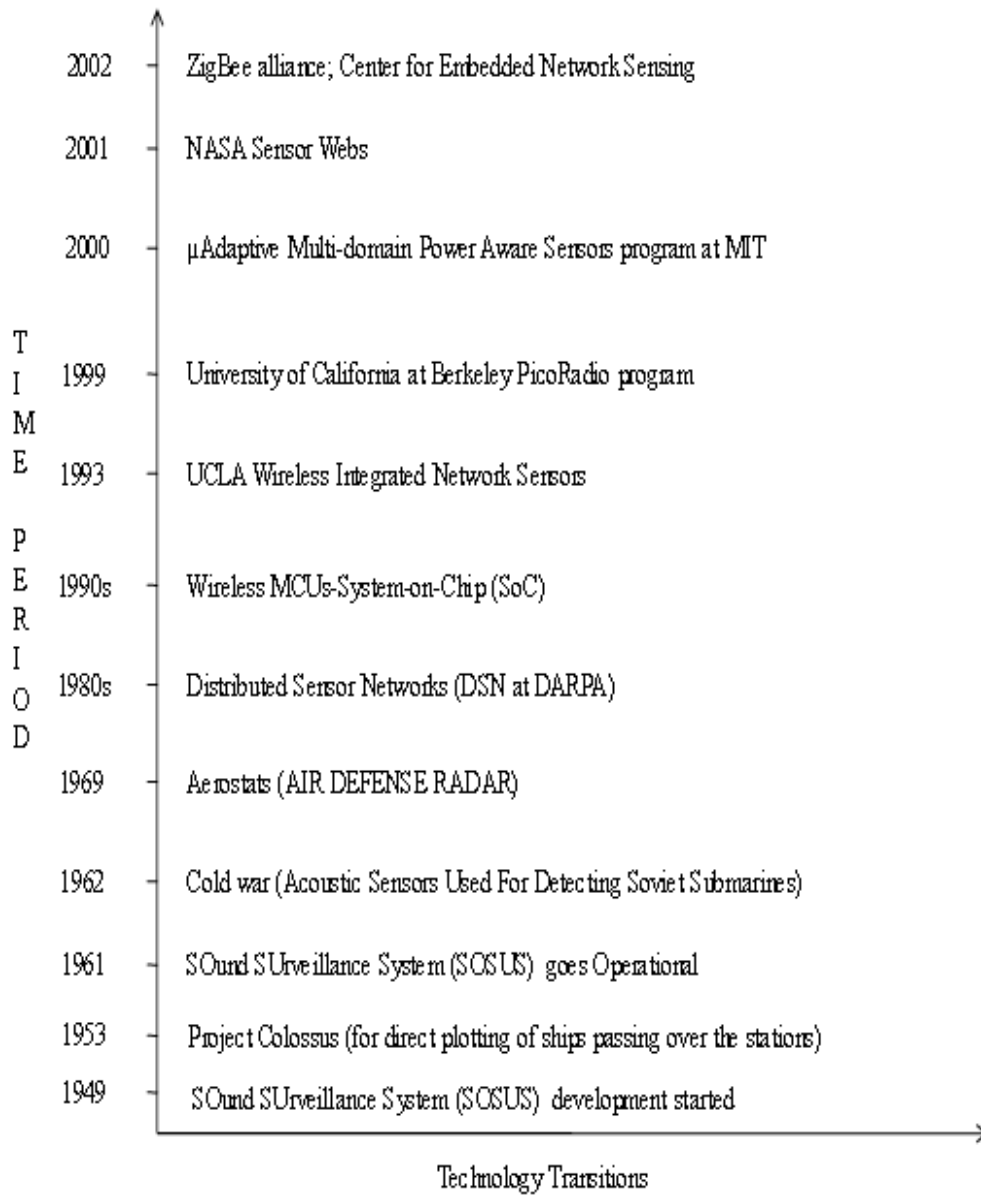


Figure 1.1: Technology transitions in WSN

Recent advances in micro fabrication technologies have made it possible to produce tiny nodes which can house multiple sensors and have reasonable processing and communication capabilities. In addition to this, development of wireless networking standards having security, stability and minimum end to end delays have led to proliferation of WSN in to the field of control and monitoring the area which was unheard of earlier. The usage of WSN is increasing exponentially due to the features such as: Scalability, Adaptability, Convenience, Mobility, Accessibility, low cost etc.

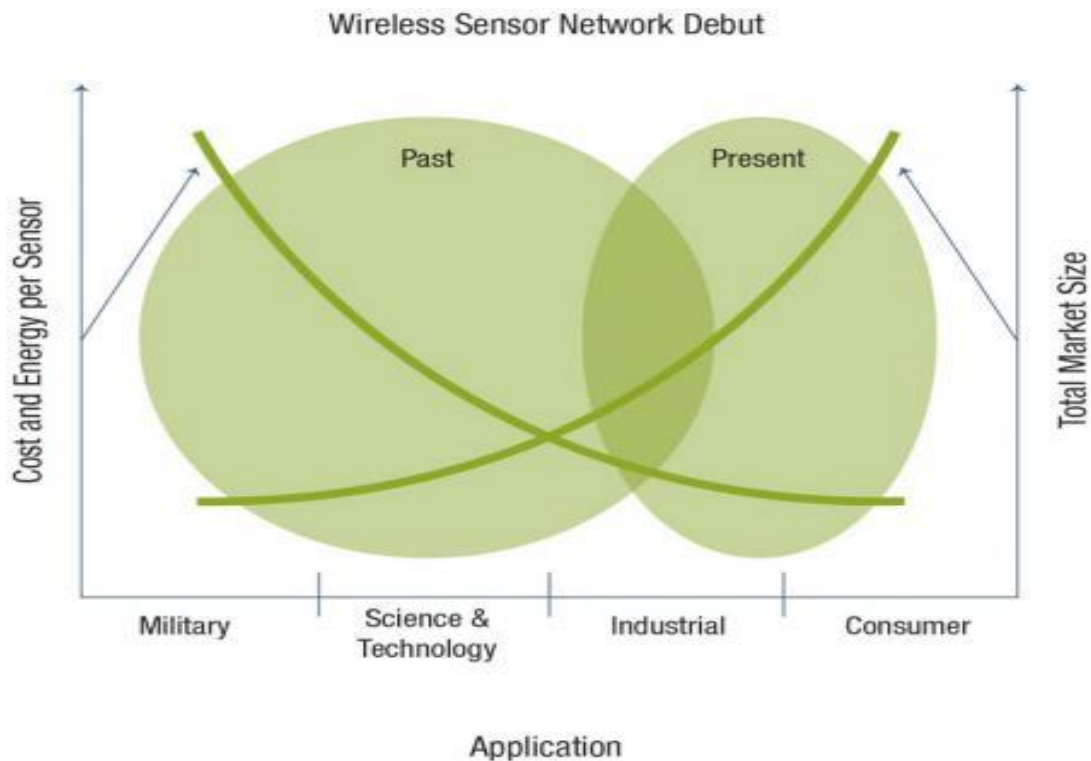


Figure 1.2: WSN gain market traction with decrease in sensor costs [104]

## 1.2 Features of WSN

The striking features in WSN are as follows:

- i) *Scalability*: WSN can be extended to any extent by adding more and more devices in random to serve more geographical area and applications. Strength of a WSN lies in numbers. As nodes are added WSN interconnections become stronger

which is contrary to other wireless networks such as - cell phone network, where increasing number of active cell phones in a small area hamper the service. Nodes are self configuring; therefore they build themselves up into a meaningful collaborative network.

- ii) *Adaptability*: These networks are able to adapt dynamically to changing environment or needs. Inbuilt easy adaptation mechanisms are quick to respond to changes in network topologies or failure of nodes. Network can shift to vastly diverse modes of operations.
- iii) *Low cost networks*: Besides low cost of nodes, no cable laying/ routing or pre-existing infrastructure is required to setup these networks. Each individual node when added to the network becomes part of the overall infrastructure called wireless sensor network.
- iv) *Convenience*: Wireless communication technologies provide freedom and distributed capabilities at affordable cost which has made their use ubiquitous. These are capable of providing spatio temporal information according to the needs/ demands of the users.

### **1.3 WSN**

Randomly deployed innumerable self configuring nodes having sensing, computation and communication abilities make a WSN. Individual capabilities of these nodes may be minimal but their build up in a network opens up vast areas of diverse application possibilities that may be real-time tracking, monitoring of environmental conditions or in-situ monitoring of health of structures, equipment or humans.

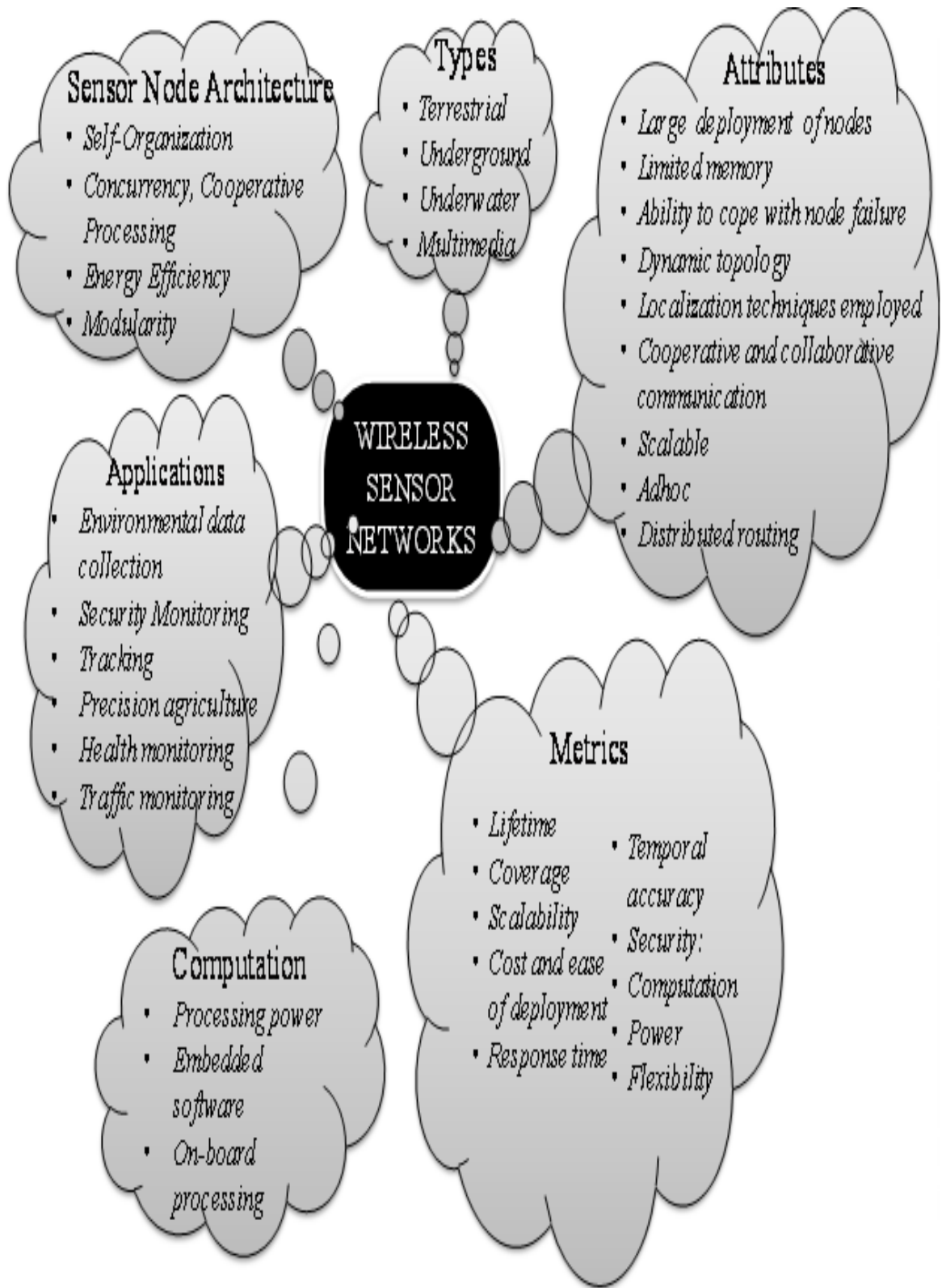


Figure 1.3: Abridgement of WSN

Development of usage of WSN was initially started from military applications like battlefield surveillance which gradually grew into the areas like applications of healthcare, home automation, traffic control etc. Advances in wireless networking, micro-fabrication & integration using Micro Electro Mechanical System (MEMS) technology and embedded chips called, micro-processors have led to the evolution of a new generation of WSN, which are capable of performing numerous applications which hitherto have been unbelievable.

Applications of WSN are going to revolutionize the way we live [1-3]. It is not a distant possibility when these sensors would be sprayed on the roads, walls, machines etc. for monitoring vehicular traffic, track job flows and supply chains in smart factories etc. Each node in a sensor network is equipped with various sensors, a radio transceiver, a microcontroller and an energy source (usually battery). Size and cost constraints on sensor nodes force the manufacturer to introduce corresponding constraints on available resources like energy, memory, computational speed and bandwidth in a node.

A sensor network is usually set up as a wireless adhoc network where each sensor supplies multi-hop routing. WSN can also operate actuators controlling various processes thus extending their control to physical world also.

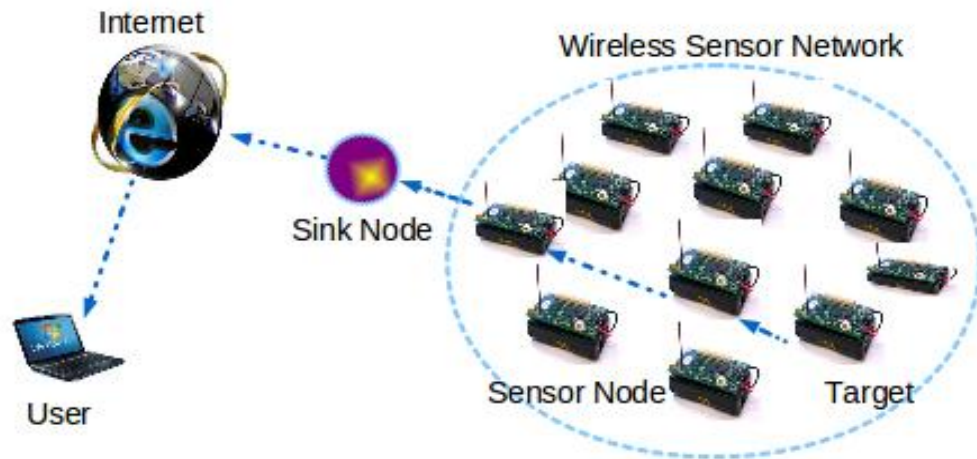


Figure 1.4: A Wireless Sensor Network [101] comprising of Sensor nodes and Base Station

### 1.3.1 Sensor Node Architecture

Dynamic random topology of WSN and collaborative processing by nodes in a distributed setup makes classical network organization concepts irrelevant. Separate programming techniques and technologies are required to cater to special characteristics of WSN which make their software architecture special and vibrant. Software architecture is unique so as to solve inherent difficulties[4] of these networks. Special characteristics influencing the software architecture are:

- i) *Self-Organization*: Due to large number of nodes in a WSN network organization [5] by manual means of users is very cumbersome. The nodes are capable of organizing themselves in a meaningful configuration and partition for efficient operation. The nodes of a sensor network are robust [6] and the aggregate formed by the nodes is required to tolerate any device failure and change in topology. It requires complex software routines to be implemented for fault tolerance.
- ii) *Concurrency, Cooperative Processing*: The nodes in a network control the information flow through the network. Each sensor node possesses a limited amount of memory, so buffering of data is impractical [6]. A node has to perform a number of operations such as: acquiring, processing and transmitting sensor data and then forwarding this data along with data received from other nodes to a gateway or cluster head depending upon network topology and routing. WSN relies on cooperative processing. Cooperative processing reduces network traffic through data aggregation and processing [5].
- iii) *Energy Efficiency*: WSN generally operates in tough and inaccessible areas therefore energy efficiency is the prime factor of importance which need minimization of communications. This has led to the development of various protocols and network configuration. Additional power savings can be given by

Operating System (OS) for the nodes by supporting advanced power management and task scheduling [5].

iv) *Modularity*: In case of special networks sensor nodes are application specific and contain the hardware needed for that special application only. The wideranges of possible applications create a large variance in hardware required for sensor nodes for varying applications. Therefore, the software for the nodes should show high degree of modularity[6].

**1.3.1.1 Network Layer Stack of WSN**

Architecture used in WSN follows generally five layers of OSIModel. In it also worth mentioning that because of multi-hop communication, the Transmission Control Protocol (TCP) is not suitable for WSN.

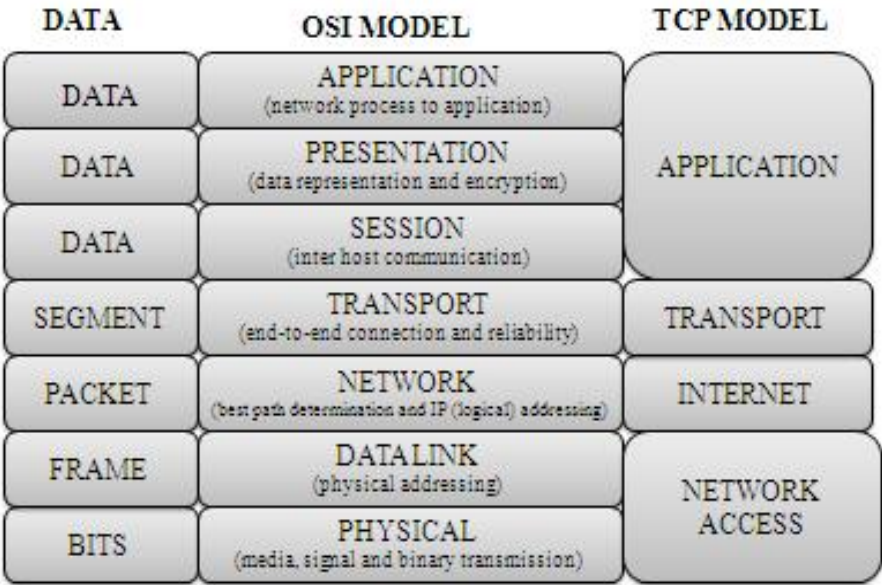


Figure 1.5: Comparison between the OSI Model and the TCP Model [102]

Different layers used in it are: application, transport, network, data link and physical. In addition to these, the special tasks of a WSN such as power management, mobility management and task scheduling to increase the efficiency of the network are controlled by three cross layers.

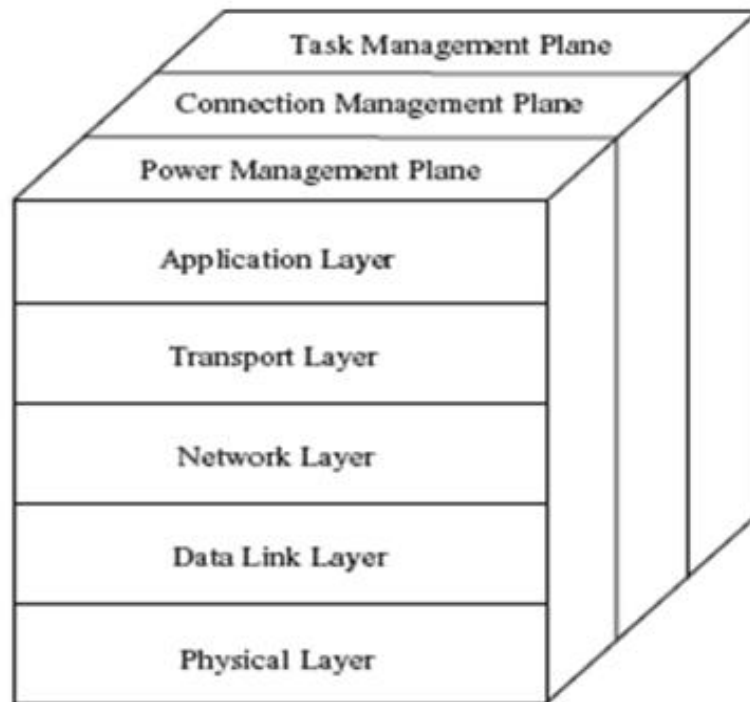


Figure 1.6: Network layer Stack of WSN[103]

- a) *Application layer*: It manages traffic and provides software for various applications which send queries to obtain information.
- b) *Transport layer*: This layer is required in case of internetwork communication. Many protocols have been designed to provide reliability and congestion avoidance. Because of multi hop communication TCP is not suitable for WSN.
- c) *Network Layer*: This layer serves the function of routing which is a challenging task in WSN. Due to low power, limited memory, routing protocol has to provide

reliable and redundant paths, for which many protocols are available according to the desired metric. This layer protocols may either be divided into flat routing or hierarchical routing or can be divided into Time driven, Query driven or Event driven. To ensure reliability in case of node failure, redundant nodes are deployed which result in production of a lot of redundant data. This data can be processed as processing consumes less power as compared to communication. It is achieved through Data Aggregation and Data Fusion. Data Aggregation is combining the data from many sensor nodes into meaningful information and eliminates redundancy. It is used in flat routing [7]. Data Fusion is further processing of aggregated data such as reduction of noise from aggregated data [7]. Some Data centric Routing Protocols are SPIN, Directed Diffusion etc.

- d) *Data link layer*: It ensures reliability from point-to-point or point-to-multipoint. Error control and multiplexing of data streams is also done by this layer. In WSN, Medium Access Control (MAC) has an important role to play. It provides higher efficiency, reliability, low delay and higher rates of communication [8].
- e) *Physical layer*: It provides an interface to transmit streams of data over a physical medium. Selection of frequency, generation of carrier frequency for modulation, signal detection and security etc. are done in this layer. IEEE 802.15.4 is one such protocol.

### **1.3.1.2 Network Components**

WSN is composed of devices (nodes/ motes) capable of sensing the environment of the area under study. The data sensed by the nodes is transmitted to the sink (Base-Station(BS)) either through single hop or through multi hops. As multi-hop saves energy so it is preferred and most widely used method of communication. These nodes may be stationary or moving, homogeneous or heterogeneous and be aware of their location or not. The diagram below shows a typical node. Main components of a wireless sensor

node are various sensors, a micro controller, memory, transceiver, and Analog-to-Digital Convertor (ADC).

- i. *Micro controller*: The micro controller operates at low frequency compared to traditional processing units. It performs various tasks such as: processing data and controlling the functionality of other components of the sensor node.
- ii. *Transceiver*: Sensor nodes operate in Industry, Scientific and Medical (ISM) Band, which allows a free radio spectrum allocation and global availability. Radio frequency-based communication is the most relevant mode which suits most of the WSN applications. WSN generally use license-free communication frequencies: 173 MHz, 433 MHz, 868 MHz and 915 MHz and 2.4 GHz. The function of both transmitter and receiver are combined into a single device which is called transceiver. The operational states of the transceiver are transmitter, receiver, idle and sleep. The transceiver may be in anyone of the states such as transmitting, receiving, idle or sleep.
- iii. *Memory*: The most relevant kind of memory is the on-chip memory of a micro-controller. Memory requirements are usually application dependent. They are employed for storing the application related data and for programming the device. Memory storage is only a few kilobytes (kB).
- iv. *Battery/Power Source*: The most critical aspect in the development of a wireless sensor node is to ensure that there is always sufficient energy available to power the system. Power is consumed by a node in sensing, communicating and data processing. The energy required in data transmission is very large as compared to any other process. Power source is generally a low power between 1.2volts to 3.7 volts battery.
- v. *Sensor*: This is a hardware device which responds to a change in a physical condition like temperature, pressure etc. and converts it into a measurable analog signal. It measures the physical parameter to be monitored. The analog signal

generated by the sensors is digitized by an ADC and sent to controller for further processing. Sensor may be of any category like passive omni-directional, passive narrow-beam sensors and active sensors. Passive omni-directional sensors sense the data without intruding the environment. They are self powered and require energy to amplify analog signals and have no means of distinguishing directions of sensed conditions in measurement. Passive narrow-beam sensors have well defined notion of direction of measurement (like a camera) and active sensors actively probe the environment e.g. SONAR or RADAR sensors.

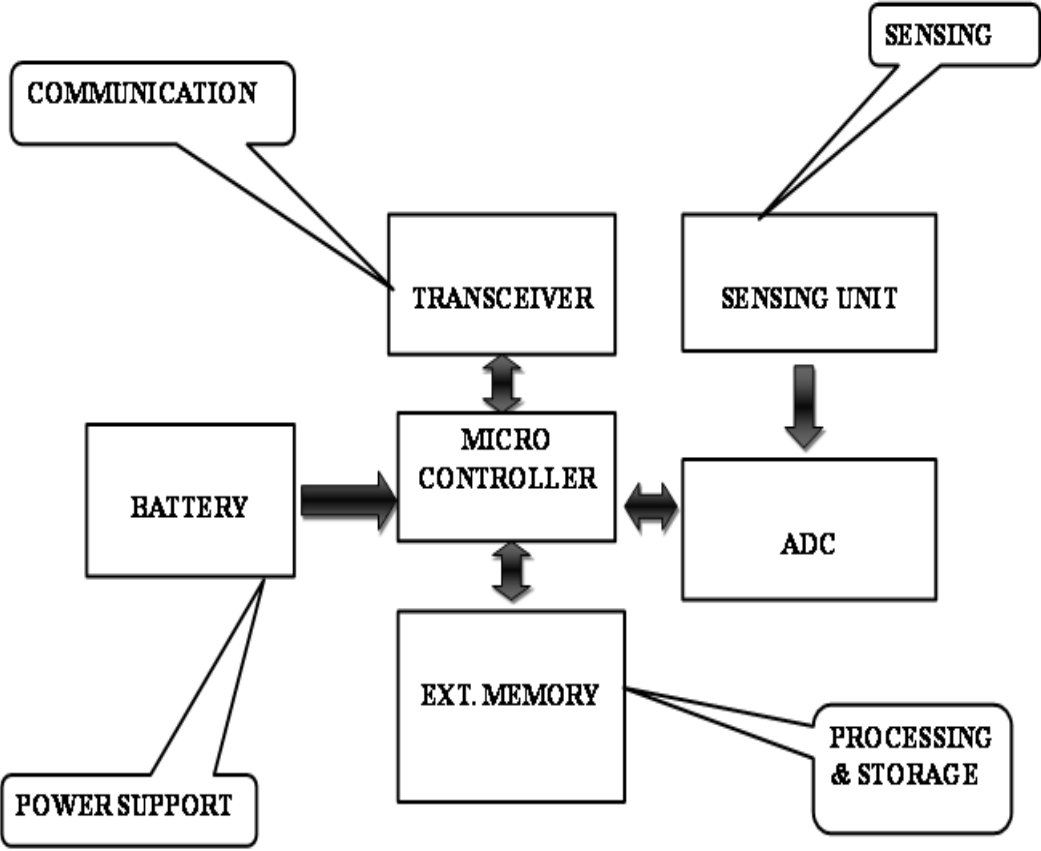


Figure 1.7: Sensor node Architecture & tasks performed by various components

Sun Small Programmable Object Technology (Sun SPOT) mote, a WSN mote developed by Sun Microsystems [9] that has been used for our experimental verification has: 180

MHz 32-bit ARM920T core processor, 512K RAM, 4MB Flash, 2.4 GHz IEEE 802.15.4 radio with integrated antenna, AT91 timer chip and a USB interface. The sensor board is equipped with 2G/6G three-axis accelerator, temperature and light sensor, 8 tri-color LEDs, 6 analog inputs, 2 momentary switches and 5 general purpose I/O pins and 4 high current output pins. It is equipped with 3.7V rechargeable 750 mAh lithium-ion battery. It has 30 $\mu$ A deep sleep mode power requirement. Unlike other available mote systems, the Sun SPOT does not use TinyOS it is built on the Squawk Java Virtual Machine.

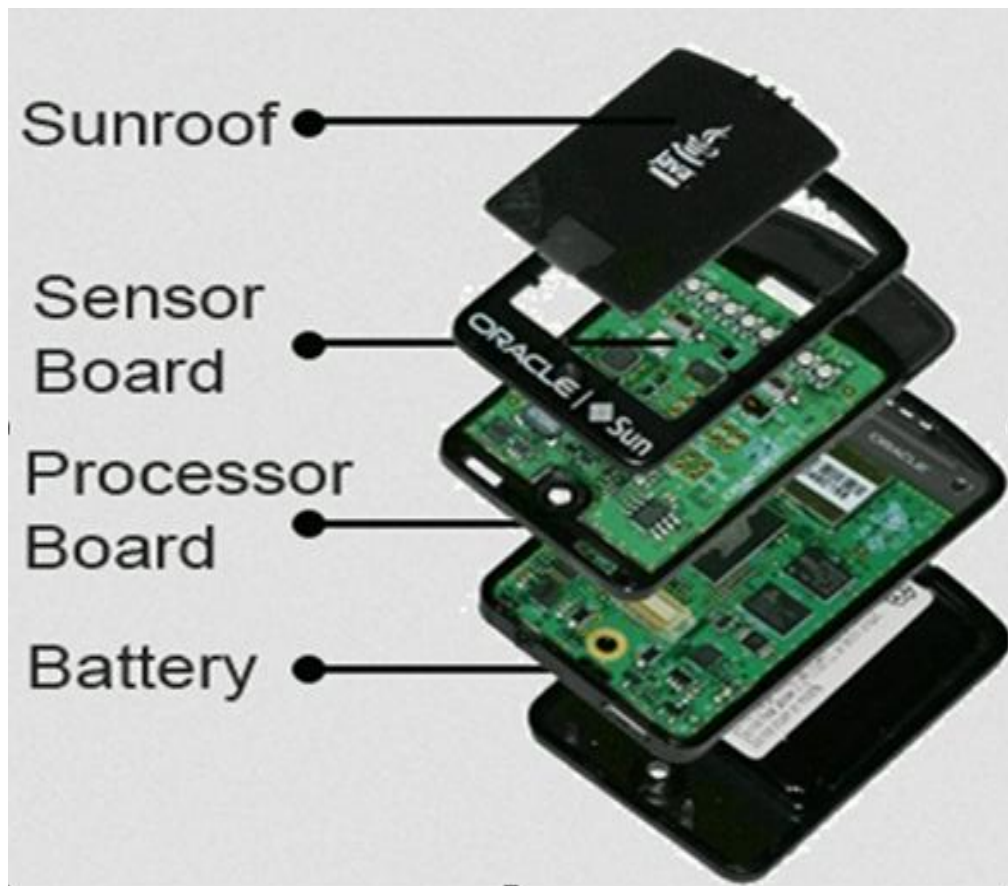


Figure 1.8: Anatomy of a SunSpot [9]

### 1.3.2 WSN Infrastructure

WSN follows ubiquitous computing [10] which is more popularly known as pervasive computing. Ubiquitous computing is an advanced computing concept where computing is made to appear everywhere and anywhere. In contrast to desktop computing, ubiquitous computing can occur using any device, in any location, and in any format. The general WSN infrastructure is made up of Mobile Hosts (MH)/ Stationary Hosts (SH) or motes and BS. The communication between the motes and the base-station is wireless whereas the communication between the base-station and the Fixed Host (FH) i.e. a PC is wired

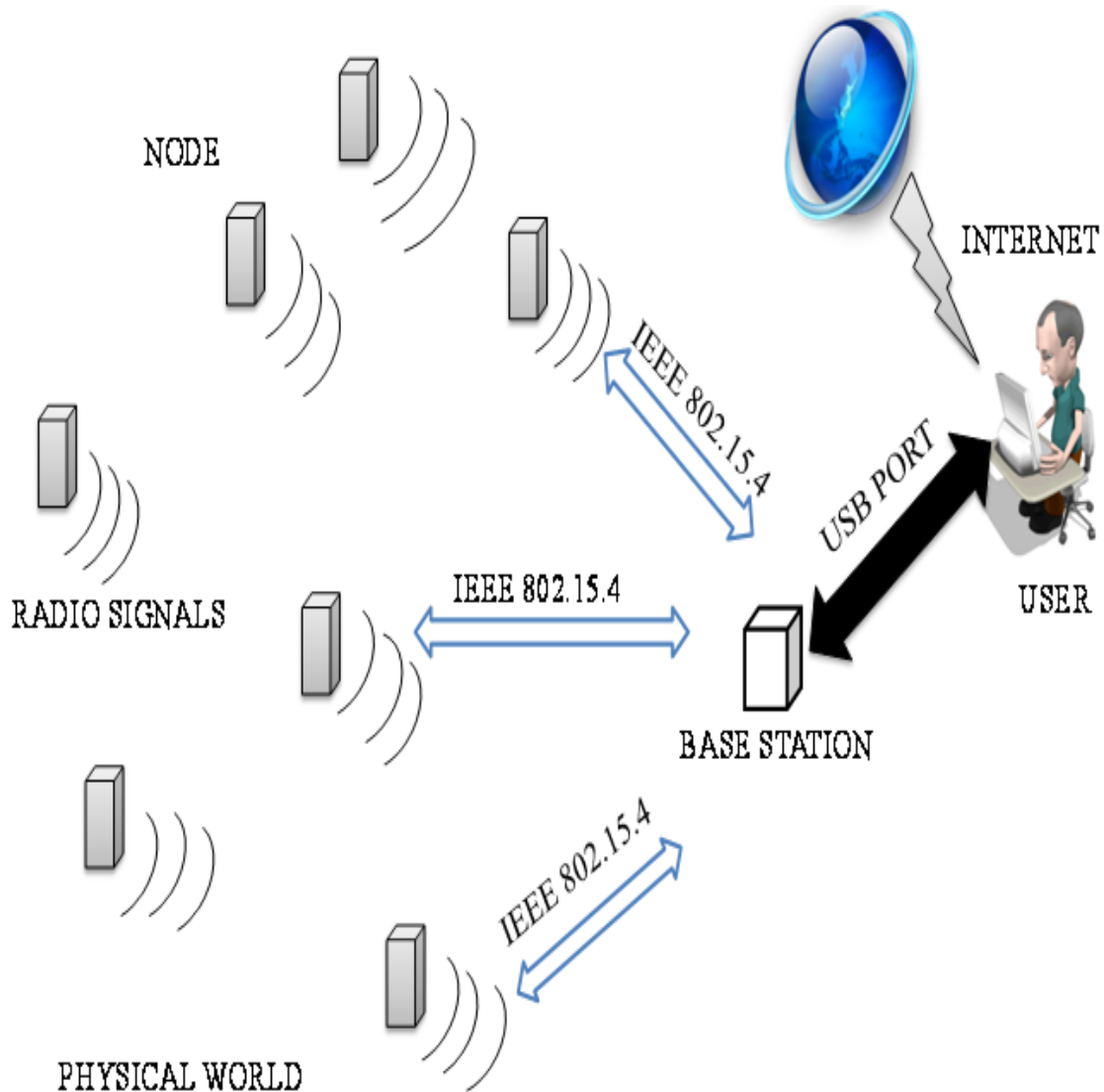


Figure 1.9: A WSN comprising of Base-Station and free Motes

### 1.3.3 Attributes of WSN

WSN does not require any infrastructure, it itself is an infrastructure which grows according to the needs and serve the objective of ubiquitous computing. WSN infrastructure has some basic characteristics which are enumerated as below:

- i. WSN may consist of thousands of nodes which are far big in number than traditional networks.
- ii. Limited amount of data is stored in these nodes due to small size of the nodes.
- iii. WSN are required to withstand harsh environmental conditions therefore are programmed and equipped to cope with node failures.
- iv. WSN are able to change topology dynamically and hence are suitable for mobile applications also.
- v. Communication in a WSN is data centric and very short packets are communicated due to low bandwidth and power. Therefore network layer overheads should be very small. This makes protocol design for a WSN a challenging task.
- vi. Communication failures are frequent due to failure of intermediate nodes from the source to the sink (BS) or due to battery exhaustion. Redundant links or nodes are deployed to maintain communication.
- vii. In order to gather information like the temperature, pressure, light etc. from the environment it is essential that the nodes are deployed in large quantity at random.
- viii. Node location is important in these networks as they are deployed for monitoring of physical environment. Therefore, various localization techniques are used in WSN.
- ix. In these networks nodes do not compete with each other for resources rather nodes work in a cooperative and collaborative manner to accomplish same goal of providing desired Quality of Service (QoS) for the designed application. As size of traditional wireless networks increases problem of spectrum congestion is faced and efficient division multiplexing techniques

are required. Sarma et al. in [11] has proposed application of Artificial Neural Network (ANN) in wireless communication where as in case of WSN no such problem is faced as sensors co-operate with each other moreover multi-hop transmission make re use of channel in the network possible so limited bandwidth can be put to optimum use.

All these characteristic unique features of a WSN make it imperative that network protocols are designed by keeping these requirements in mind and which can use constrained resources efficiently. However, packing all these features in a single protocol is not feasible in multi-hop small packet communication. Therefore a number of applications specific protocols from data link layer to transport layer have been devised which suits best to the requirements of the application.

Atul M. Gonsaiet al. in [12] have described network routing algorithms for WSN which can cope with the problems of path failures, packet loss, dynamic nature of topology and node density. JianPing Wang et al. in [13] have proposed use of nature inspired algorithms swarm intelligence and Ant Colony Optimization (ACO) for developing routing algorithms for WSN. Menik Tissera et al. in [14] has proposed a new method for information extraction from WSN with minimum latency and energy usage thereby making them suitable for mission critical applications.

### **1.3.4 Applications of WSN**

Sensor network applications are so numerous and varied that volumes are required to enumerate even some of the major ones. However, to have an idea of these applications, the applications may be classified broadly into three classes: i) environmental data collection ii) real time process or security monitoring and iii) tracking.

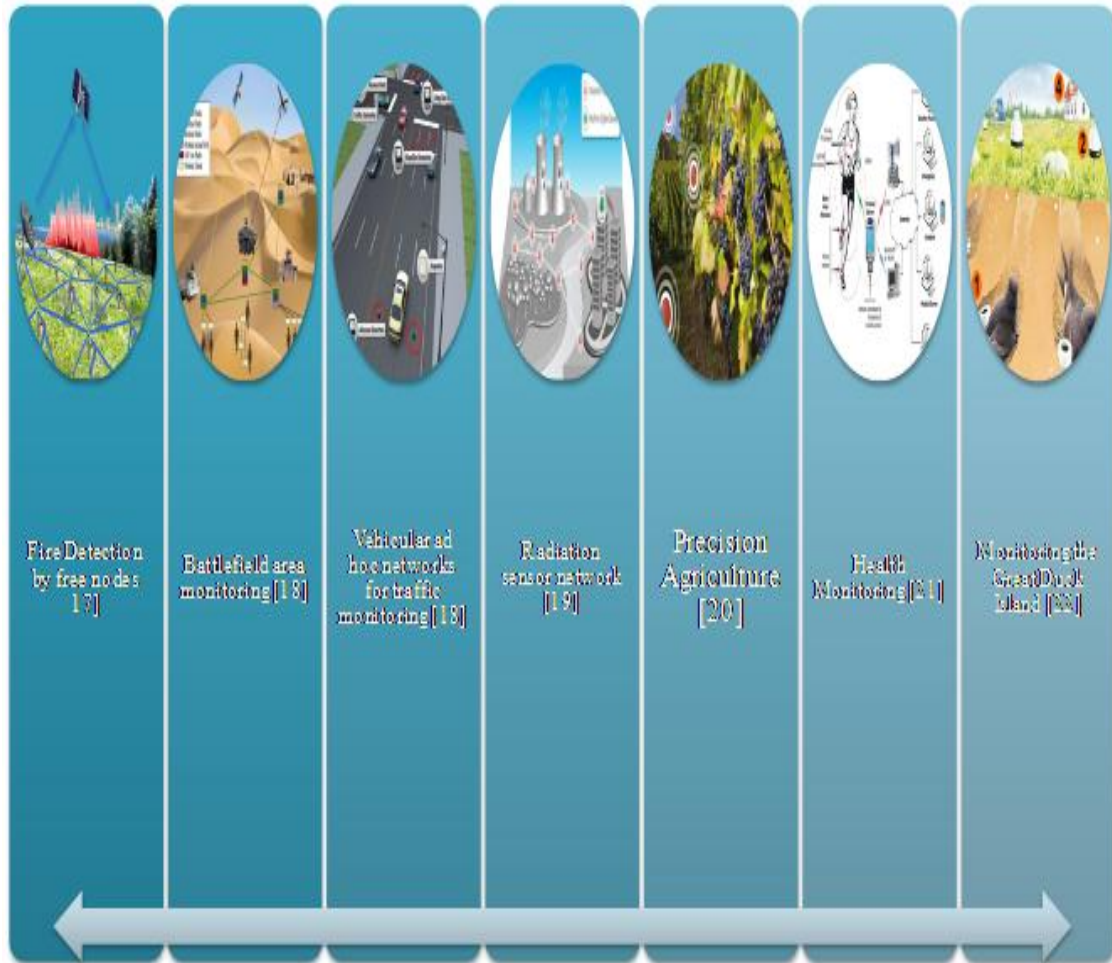


Figure 1.10: Applications of WSN

- I. *Environmental data collection*: Basic feature of such networks is a collection of large number of nodes which require very low data rates and extremely long lifetimes. The nodes are evenly distributed over an environmental area of study. Nodes are spread at minimal distance to cover the area fully so that no details are missed out. The nodes used are capable of arranging themselves in a reliable topology and develop optimal routing strategies which can be used to route data to the central collection points. It is not essential that the nodes develop optimal routing strategies on their own. Possible optimal routing topology can be calculated outside the network and then communicated to the nodes. This is feasible as the physical topology of such networks remain static over long duration of time therefore routing topology once transmitted is valid for long

periods. In environmental data collection applications tree-based routing topologies are used. Here root of each routing tree is setup at high-capability node which sends the data to the sink. Data is transmitted from the child nodes to the parent node upwards in the tree till it reaches the sink. In tree based data collection, each node is responsible for forwarding the data of all its children. Nodes having large number of children have to transmit significantly more data than the leaf nodes. These nodes can quickly deplete their energy therefore may require change in topology to assign more responsibility to a node having more energy so that node failure can be avoided. Once the network is configured, each node periodically samples its sensors and transmits its data up the routing tree and back to the BS (sink). The typical environmental parameters like temperature, light intensity and humidity do not change quickly therefore sampling rates may be more than 15 minutes. Besides low sampling frequency, environmental monitoring applications do not require quick response as data is collected for the future analysis and there is no quick change.

We can say in order to meet high lifetime requirements duty cycle is to be optimized. The network is to be reconfigured to handle node/ link failure by redistributing network load. In addition to this, as the researchers learn more about the environment under study they may require finer details and insert additional sensing points. In both the cases the reconfigurations are required but these are relatively infrequent and therefore do not use much energy.

The most important characteristics of the Environmental monitoring networks are long-life, precise synchronization, low data rates and relatively static topologies. As real-time data transmission is not critical, the data transmission can be delayed inside the network if it can improve network efficiency.

- II. *Security Monitoring*: For these applications sensor network is generally not required to collect or transmit any data at a predefined rate. Therefore, optimal network architecture is such that each node has to do self diagnosis and check the

status of other nodes assigned to it at regular intervals. Responsibility of status checking of surrounding nodes is evenly distributed among some nodes in the network. Any security violation alarm is to be reported without any delay in a deterministic way. This requirement, needs all the nodes to be kept active all the time which results in unnecessary energy consumption during idle period. In these networks, transmissions are rare therefore transmission costs are not much significant as compared to costs involved in ensuring minimum latency.

These networks are made up of nodes placed at fixed locations throughout the environment which is to be monitored continuously. The main difference between these and environmental monitoring networks is that these networks actually do not collect any data. This fact has a significant impact on arriving at the optimal network architecture. Each node has to check the status of sensors but it has to transmit a report only when it observes any deviation from the setup. The immediate and reliable communication of alarm messages is the primary requirement of the system. These networks are known as “*report by exception*” networks. It is also essential that each node must confirm its status or may be its neighbor also. The optimal topology of a security monitoring network may differ vastly from that of a data collection network.

- III. *Tracking*: In this application object to be tracked is tagged with a sensor node. It is tracked as it moves through a WSN area covered with sensor nodes which have been kept at known *locations*. Sensor nodes of WSN in this case are to sense RF messages from the node tagged to the object under observation. These networks are required to sense continuously moving objects which are leaving or entering the network.

The WSN application may also be classified as i.e. Event Detection (ED) and Spatial Process Estimation (SPE). In ED the sensors are deployed to detect an event e.g. fire in a forest, a quake etc. The number of nodes deployed depends upon the area to be covered and communication range. In SPE the main task is to

get estimation of the behavior of the spatial process based on samples taken by the sensors. Environmental monitoring is an example of application belonging to both the categories. Variety of possible applications of WSN to the real world is practically unlimited. Today we have entered the 3<sup>rd</sup> wireless revolution, “Internet of things”. The 3<sup>rd</sup> wave uses the wireless sense & control technology to bridge the gap between the physical world of humans and the virtual world of electronics.

## 1.4 Types of WSN

Depending upon the environment or area of operation the WSN may be categorized as follows:

- I. *Terrestrial WSN*: On ground networks, these networks can either be Ad-Hoc (unstructured) or Preplanned (structured) in nature.
- II. *Underground WSN*: These usually fall into the category of preplanned i.e. structured network. The equipment involved is very expensive, and after deployment, regular maintenance is required.

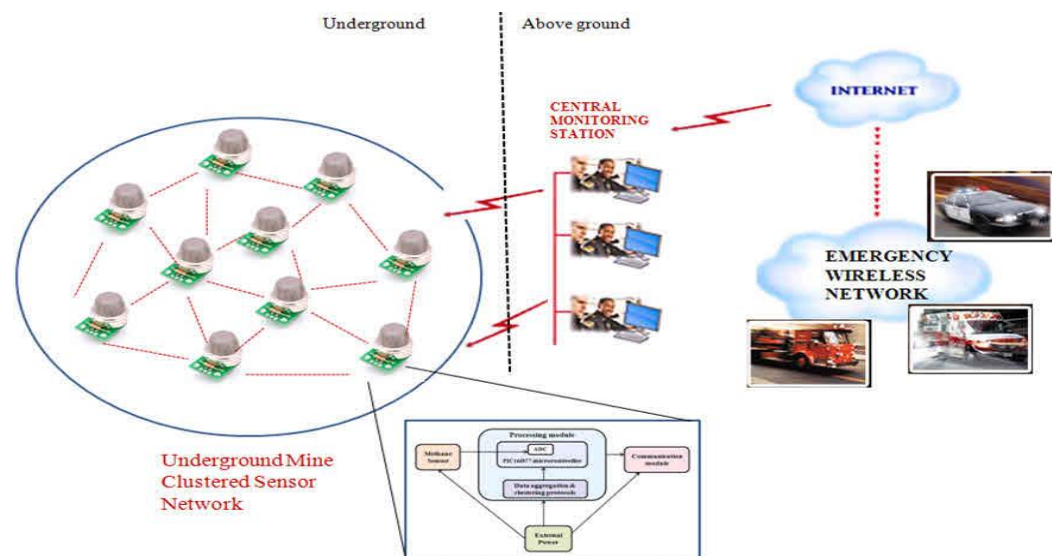


Figure 1.11: Underground WSN[105]

- III. *Underwater WSN*: The sensor nodes are sparsely deployed. The network is more expensive than the terrestrial WSN. The technique adopted by the WSN is underwater acoustic wave communication. Under water communication is difficult due to small available bandwidth and long propagation delay.

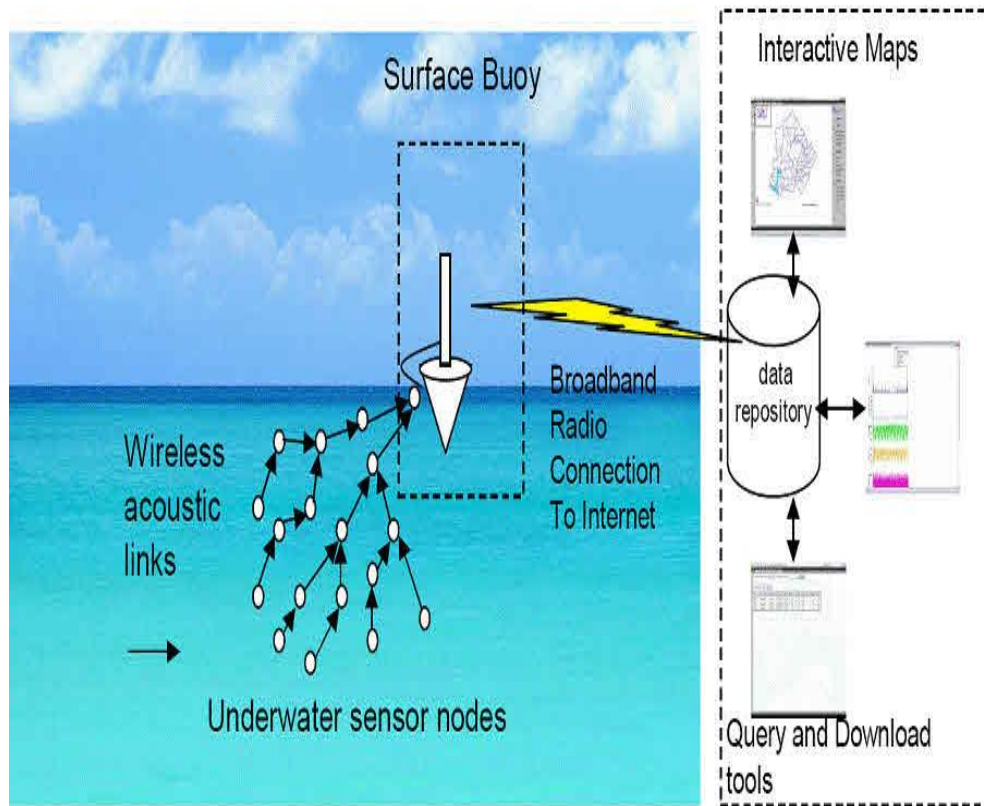


Figure 1.12: Underwater WSN [105]

- IV. *Multi-media WSN*: These employ application specific devices. The network is equipped with cameras and microphones. High bandwidth/ low energy, filtering, data processing and compressing techniques are adopted for execution of the applications and deliver required QoS.

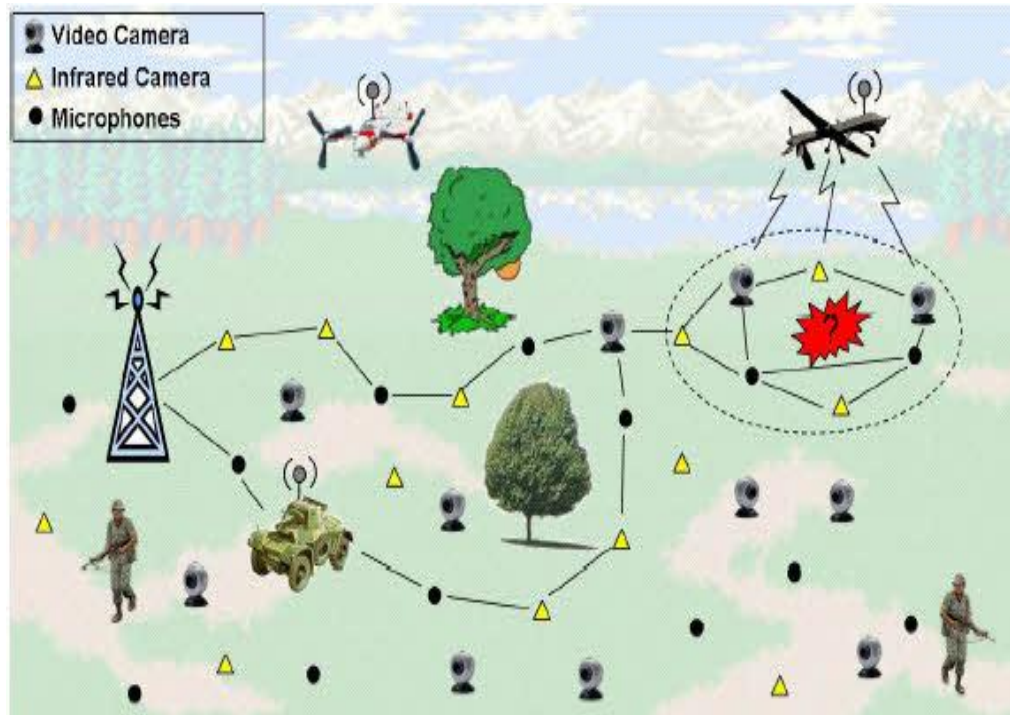


Figure 1.13: Multimedia WSN [105]

- V. *Mobile WSN*: These networks may be defined as WSN where the sensor nodes are mobile. The nodes are versatile and capable of repositioning and organizing into the network. Initially the sensors are deployed and these then spread out to gather information.

## 1.5 The Underlying technology

As discussed above, various application specific communication protocols are being employed by the users to gather required information with the desired QoS. However, some sort of standardization is required to make the things interoperable globally. In the area of WSN, various standards are either waiting for a formal acceptance or are under development[15]. The major standardization bodies dedicated towards standardization are Institute of Electrical and Electronics Engineers (IEEE), the Internet Engineering

Task Force (IETF), the International Society for Automation (ISA) and the HART Communication Foundation etc. Various protocols established by IEEE to suit the varied requirements of range, heterogeneity of devices in the networks, cost, security and fidelity of data are covered under various IEEE 802 standards which have evolved over time. 802.11 series of standards governs the communication between wireless networks. However, to support newer modulation techniques for higher data rates and low susceptibility to interference 802.15.4 series standards have come up.

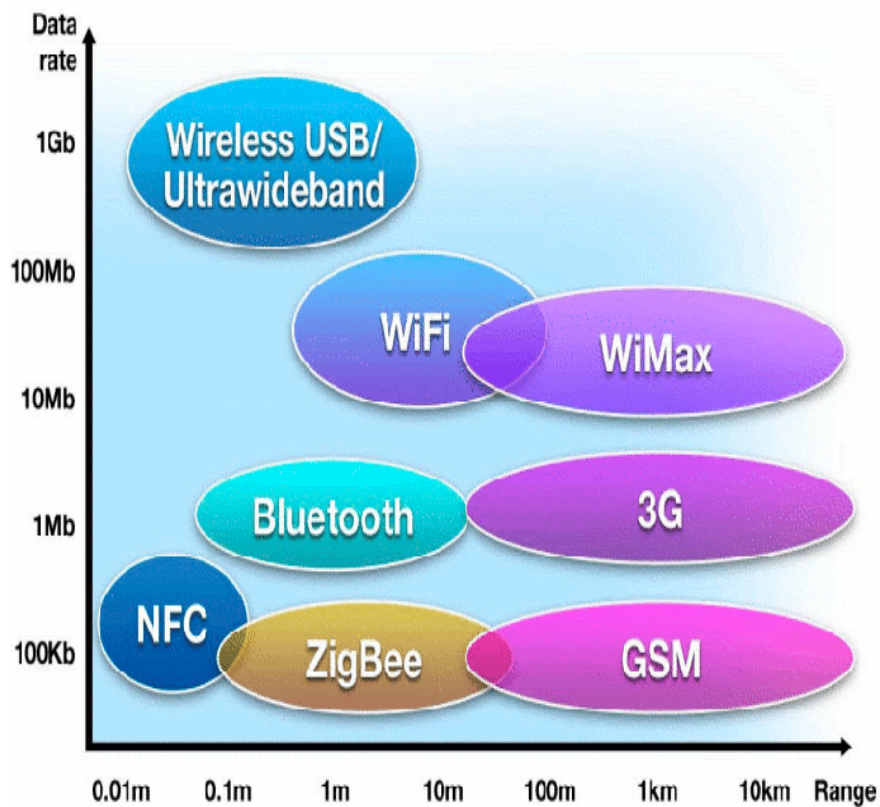


Figure 1.14: Wireless sensor communication technologies comparing data rate and range[106]

The first standard for WSN, based on IEEE 802.15.4 was ZigBee. It was developed with a target to make it a common standard for industrial, commercial, residential and many more applications. ZigBee[16]hasupper layer specification built upon IEEE 802.15.4. It is a standard employed by Wireless Personal Area Network (WPAN) functioning in the 868/915 MHz and 2.4 GHz ISM bands along with Bluetooth [17]. However, according to

[18], ZigBee is highly unsuitable for industrial applications as it does not comply with the stringent requirements of industry. It provides no security/ reliability and does not guarantee the maximum latency in end-to-end communication delay required by monitoring applications that need to communicate with sensor nodes within a second. As new applications of WSN are envisaged in industry also, further research led to development of many more standards. These are based on ISA 100.11a. WirelessHART is devised for process monitoring applications.

## 1.6 Query processing

In a WSN, huge amount of data is generated by sensor nodes continuously. Data generated by nodes make a large distributed database. Information requirements for various applications can only be satisfied if the data available in the network is extracted to a suitable location and processed efficiently. However, query based systems are popular in database type environment. For extraction of information from the data, queries are injected as in other databases. A good DBMS is a system which optimizes on resource consumption in response to a query. Query processing is an elementary part of any DBMS. The means by which we can obtain the best plan, used in implementing the data extraction request from database is “*Query Processing*”.

Huge repository of data is available in the form of various databases. To make this data into information, it is essential that the user be able to extract precise and relevant data with minimum consumption of invaluable resources- time and energy in response to the query injected by the user. Query processing is an elementary part of any DBMS. Query is a tool to gather information. The word derives from the Latin “*quaere*” i.e., to ask or seek.

We can choose parameters from the list called menu which contains all the available parameters in the database. Various manipulations like retrieval of available information from the database, insertion of new information into the database, deletion of information from the database and modification of the stored data are possible through Data

Manipulation Language (DML). Query language is a subset of DML which is used for writing a query. Query processing is the means to obtain the best plan for executing the database request.

### 1.6.1 Query Types

Various types of queries are described below:

- I. *Select queries:* It helps in retrieving the data from various tables of the database using primary key(s). Results are displayed in a single table.
- II. *Parameter queries:* These are the queries created by the user on the basis of criteria which may be available during the running of the query. These are called on-the-fly queries also.
- III. *Crosstab queries:* These are used to generate reports using an attribute/ field. The report is then tabulated.
- IV. *Action queries:* These are used to manipulate the data of the database. There are four types of action queries:
  - i) Delete queries – The ones which are used to remove records from the tables.
  - ii) Update queries – The ones that make global changes in a group of records in a table.
  - iii) Append queries – The ones that are used to add records from one or more tables at the end of table(s).
  - iv) Make table queries – The ones that are used to create a new table by using a complete or a part of the data of an existing table.
- V. *SQL queries:* Query created by using Structured Query Language (SQL). SQL is an advanced query language.
- VI. *Contextual queries:* These may be any one of the following:
  - i) *Search context* – It is used for accessing or retrieving information from database having bulk of data such as Web Indices, bibliography, catalog etc. The contextual query language may have a single or multiple search clause(s) joined by Boolean operators. They are also

associated with keywords, which may either be at the start or at the end of clause.

ii) *Taxonomy* – The parameters are represented in the form of trees having a root and various nodes (children). Each node is associated with an established groups or classifications. These groups are specified by the labels. These labels run from the root to its children.

iii) *Level-N category, ancestor category and sibling category*: This is used for a particular classification (category), specifying the level taking into consideration the location (say n) of the node.

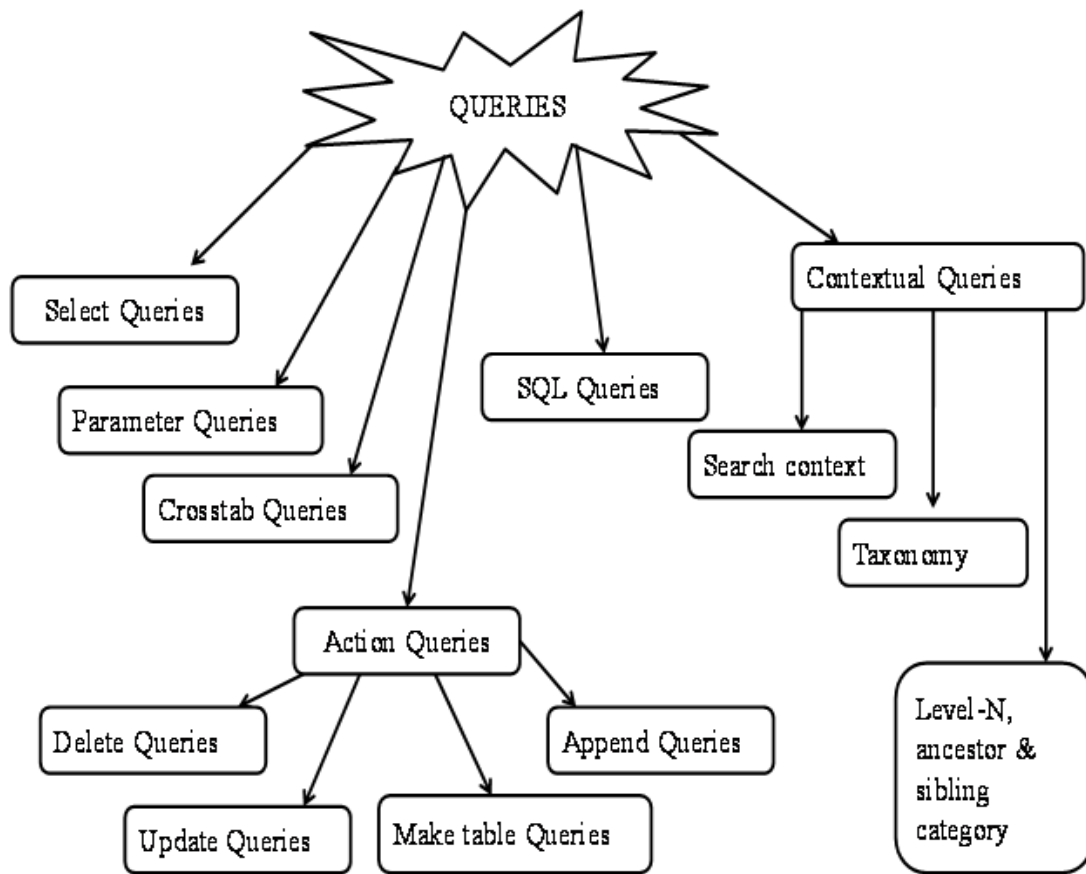


Figure 1.15: Various query types

## **1.6.2 General Strategy for Query Processing**

Generally queries posed by the user are not in a standard form. Therefore it has to be converted into a form which is understandable to digital processing systems. The query processor [19] performs the job of transforming the query into a standard form like relational calculus, relational algebra, object graph, operator graph or tableau. The query processing involves three stages: parsing and translation, optimization and evaluation.

## **1.6.3 Relationship between Query processing and WSN**

Recent studies have shown a great attention being focused on mobile computing, with intent of accessing data at any place, at any time. The devices so mentioned are not only capable of transmitting voice but are able to transmit the data as well. The main purpose of WSN is to collect data and aggregate the data. The presence of full-fledged computing hardware in the sensor node capable of performing multiple sensing functions like the temperature, light, humidity, etc., these sensor networks (which are composed of thousands of such nodes) may be thought of as a '*distributed database*'.

This database accumulates physical measurements, inside the memory available on the board about the environment and aids in providing answers to the queries injected by the users. Each sensor node produces data packets that are acquired by the sensing devices present in the node.

The nodes which are composed of micro processor, transceiver, battery, memory, sensing device and ADC enable the node to sense, compute and communicate the data. The nodes in the WSN generate the data which may be extracted by its name, which depicts similarity to the traditional view of relational databases.

### **1.6.4 Query Processing in WSN**

Unlike query processing in traditional databases, multi query processing in WSN is different because of its own semantics (of queries), objectives and resource constraints (of the node).

For the ease of users, WSN may be considered similar to a database [20-22]. This assumption makes it more comfortable for the usage along with the ease of modifiability and scalability resulting into better performance. All these factors enable WSN to execute various tasks at the same time. Considering WSN to be a database approach, query processing in WSN is executed as follows:

Initially, a user injects the query of his interest to the BS. The BS generates the query plan and communicates the query to the free nodes in the network. On receiving the query the nodes collect the required data (of the parameters specified) from the environment followed by data processing. Then the data is transmitted to the BS jointly with the other free nodes. On receiving the data final processing takes place at the BS and finally the result is submitted to the user's query.

### **1.6.5 Multiple Query Optimizations (MQO) in WSN**

Queries in WSN are continuous and we consider data as stream. Technically, an appropriate term given to it is "Window" i.e. data coming during a specific period of time as a unit. The query optimization methods may be categorized as: *Base station optimization* and *In-Network optimization*.

*Base Station Optimization:* This method helps in reduction in energy consumption in the network by reducing the communication messages as the number of injected queries in the network is reduced at the BS. This is done by finding out whether any new query submitted to the BS can be rewritten using the already running queries. This newly rewritten query is tried to be evaluated at the BS using the results of the currently running queries.

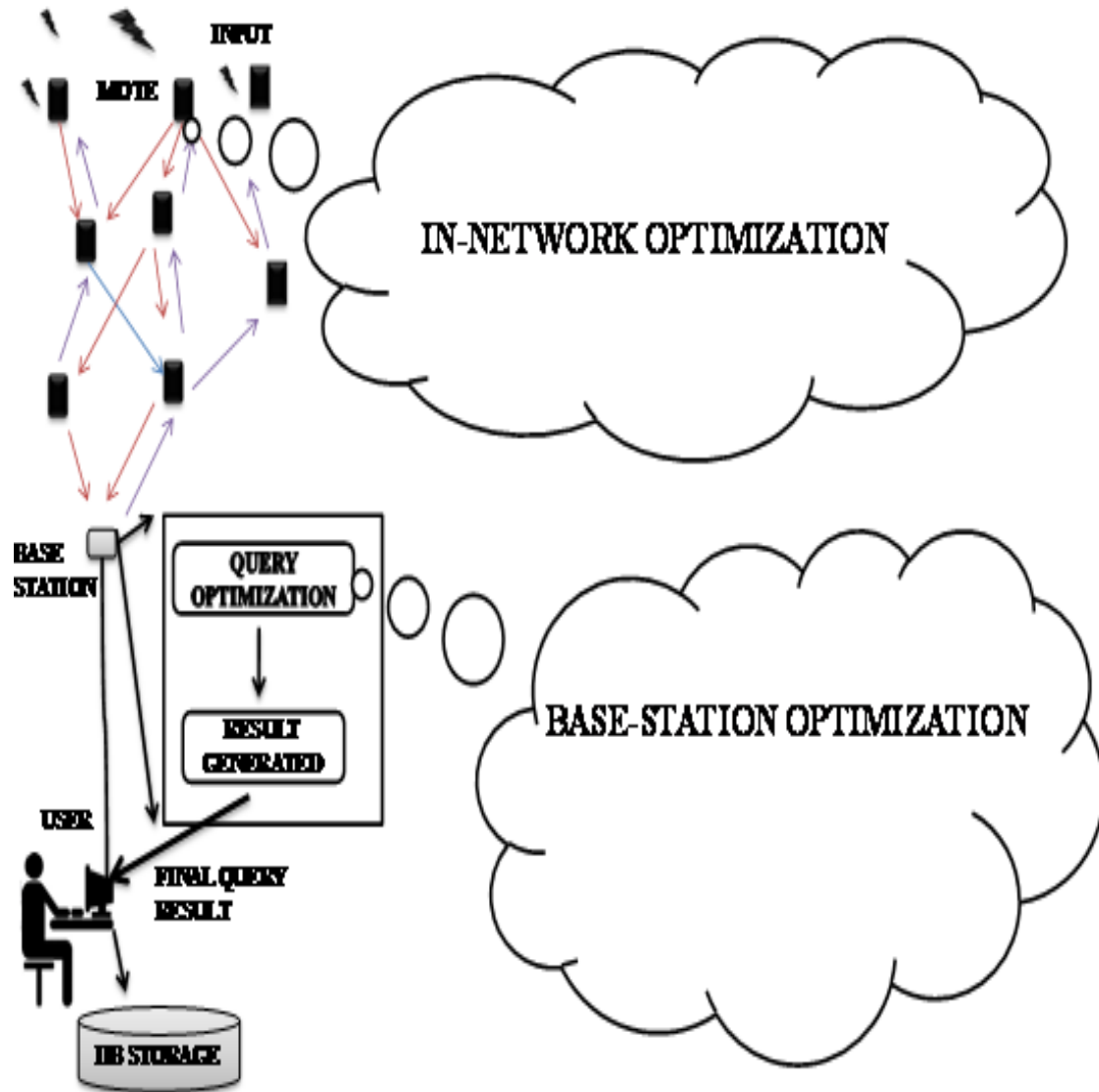


Figure 1.16: Categories of Query optimization methods (Base station optimization and In-Network optimization)

*In-Network Optimization:* The sensor nodes collect the data for all the queries that have been injected into the WSN. It is processed for deletion of redundancy etc. Resultant data is broadcasted thereby minimizing the number and quantity of radio messages thus making it bandwidth and energy efficient. Sensor nodes being resource constrained devices, it is required that the queries be processed in an efficient way [9]. Therefore, it is necessary that all the nodes should be at the disposal for supplying continuous data streams of the data gathered to satisfy the query in hand.

Various new proposals have come up regarding Two-Tier Multiple Query Optimization (TTMQO) leading to decreased number of radio transmissions and lowering the transmission time in the sensor network.

### **1.6.5.1 Key requirements of MQO**

The key requirements of MQO may be categorized as:

- i. *Scalability:* it is required that the WSN should comprehend the users' requirements by providing accurate results without getting hanged.
- ii. *Energy Efficiency:* Seeing the current scenario as the number of applications increases, energy and bandwidth would always be constrained. In order to circumvent with this problem it is essential to have optimized plan for sensor nodes.
- iii. *Adaptability:* WSN executing multiple queries should accommodate the queries according to the current workload and network condition.\
- iv. *Simplicity:* Taking into consideration various constrained resources (storage, bandwidth etc.) it is required that light weight processes be generated.

## 1.7 Metrics for WSN Evaluation

WSN find usage in a wide range of applications. It is possible due to various mechanisms such as adhoc routing protocols, in network data aggregation and signal processing etc. performance of the network can be improved by tweaking these features. In order to evaluate performance of a network objectively a number of metrics have been defined.

Some important ones are: lifetime, coverage, cost and ease of deployment, response time, temporal accuracy, security and effective sample rate. Many of these metrics are interrelated. Generally, performance in one metric may have to be decreased to achieve increase in some other metric. All these metrics form a multidimensional space where every possible application is represented by a single point. The single point represents value of all these parameters which may be high in one metric and low in other as required by that specific application. A network will be able to serve the application if that point lies in the multidimensional space of performance metric. In other words a tradeoff is to be established.

- i. *Lifetime*: One of the requirements of a WSN serving any application is to place nodes in the network which can serve for years without any intervention. Primary limiting factor in achieving this is energy supply. Efficient energy management is necessary to score higher in this metric. Significant factor to be monitored or controlled is amount of radio communication. Any gain in this front may be at the cost of other system metrics.
- ii. *Coverage*: Importance of this metric is next to lifetime only. A WSN is supposed to cover a large physical area. Multi-hop communication is used in these networks allow the coverage area to be extended independent of range of wireless communication links being used in the network. However, coverage is limited by the fact that multi hop networking protocols increase power consumption of the nodes by increasing overhead transmission costs.

- iii. *Scalability*: it is a key component determining range or coverage of the network. A fully scalable network gives confidence to the user that the network will save every eventual need of the user.
- iv. *Cost and ease of deployment*: Low cost and ease of deployment are the advantages of WSN. Ideally nodes of WSN may be placed by an untrained person randomly and the system should work. However, in actual conditions there are some constraints which should be assessed and indicated by the nodes. Therefore, each node must have adaptability to discover links and determine their quality. To have a low lifetime, network should be able to adapt to changing environment. System should be robust and perform continuous self diagnostic and generate maintenance requests if required. These features require some energy allocation for this purpose only which designing the network which may ultimately affect lifetime or it can be compromised with decrease in sampling rate.
- v. *Response time*: Response time gains significant importance in mission critical real time applications. Nodes must be capable of sending high priority messages immediately. Low response time conflicts with requirement of longer lifetime. Response time can be improved by including externally powered nodes which can report alarm messages instantaneously but this will reduce metric of ease of deployment.
- vi. *Temporal accuracy*: in many applications samples from various nodes must be accurately time stamped so as to correlate data of various nodes placed at various locations. It requires global time synchronization mechanism which requires power and bandwidth.
- vii. *Security*: Data dissemination from the network should be secure and should reach authentic user only. User of encryption and authentication measures, costs power and bandwidth.
- viii. *Power*: Each node should use power economically. This is achieved by using low power consuming hardware and optimization of duty cycle. During active

operation radio communication uses a significant portion of total energy. Application specific algorithms and protocols are developed and emphasis is to use localized computation.

- ix. *Flexibility*: Nodes should be designed with sufficient modularity to assemble right set of software and hardware suited to specific application.
- x. *Robust*: System should be robust to node failures and external interferences. Immunity from external interference can be increased through use of multi channel and spread spectrum communication.
- xi. *Communication*: Transmission range of individual nodes may not affect the coverage area of a WSN however it has a significant impact on minimal acceptable node density. Radio communication range may require higher node density than deployed to meet requirement of a particular application.
- xii. *Computation*: A wireless sensor node is required to perform “In-network data processing” and management of the low level wireless communication protocols. There are strict real-time requirements associated with both communication and sensing.

## 1.8 Gaps

Recent years have seen proliferation of WSN into almost every field where modern technology is being used. WSN are called to serve numerous applications. Applications are becoming diverse and more demanding. This has brought about the use of heterogeneous sensors, different networking topologies, protocols, and communication means and data extraction techniques in this field. Due to diversity of sensor applications evaluation and selection of a particular set of hardware setup has become very difficult as a lot of subjectivity has cropped up in performance in evaluation of WSN. In order to have evaluation objectivity, a number of metrics as described above in the chapter have

been devised which have made the job a bit easier. The research literature reviewed has shown that most of the research has been for application specific WSN where main emphasis is on a particular parameter of concern. This limits the use of WSN to that very application only. Research in communication technologies and development of various software has made it possible to use same WSN for varying application simultaneously or at different times.

However, the research work has targeted only a few metrics of the lot. In addition to this, most of the proposed techniques for communication reduction are based on data prediction or modeling which introduce some set of approximation.

## **1.9 Problem formulation**

In view of these gaps a generalized approach has been proposed which will be applicable to every type of application. There shall be no compromise on data fidelity. Researchers can further add on their application specific tools of routing, protocol selection etc. to reap more gains. Therefore, the problem for research work has been formulated as to devise a framework which is beneficial to all WSN irrespective of their application. Solution to problem of some data loss which may prove to be critical at time has also been given importance in the work. Out of all the means of data extraction query based approach is the most popular approach for data extraction. Therefore, for finding solution to the problem and popularity of query based approach following objectives have been set up for the work.

## **1.10 Research Objectives**

The objectives of research are listed below:

- I. Review of the literature on Query processing and WSN.

- II. To propose an efficient framework for Query Processing in WSN environment.
- III. To design and develop the proposed framework.
- IV. Verify and validate the proposed framework. For e.g. by using synthetic data set.

## 1.11 Thesis Organization

The contents of the thesis are summarized below:

The thesis is divided into six chapters. Concise information regarding each chapter is as follows:

*Chapter 1:* Introduction to WSN, concept of query processing, relationship between query processing and WSN, query processing in WSN, various performance metrics etc. are discussed.

*Chapter 2:* Research literature available in the field of WSN development has been reviewed.

*Chapter 3:* It gives details of our work. The objective of our work is to propose a “Framework for Efficient Query Processing in WSN”. The first part aims at processing multiple queries fired from the base station with a view to reducing energy consumption inside the WSN. In the work, an energy-efficient multiple query optimization method to reduce the number of monitoring queries running in the sensor network has been proposed. When a new query is submitted to the BS, the proposed scheme checks whether the same can be merged with the already running queries or running queries can be rewritten to satisfy the new query. As a result, the number of queries injected into the sensor network is reduced, resulting in lower energy consumption. Secondly, we also propose an In-network optimization technique leading to reduction in the data to be transmitted, leading to energy saving.

*Chapter 4:* Analysis and comparisons of the various algorithms used in the study and results of work are discussed.

*Chapter 5:* This is the concluding chapter, conclusions drawn and future scope of the research work carried out is outlined in this chapter.

## **Chapter-2**

### **Literature Review**

Recent unprecedented advances in fabrication and computing technologies have made it possible to produce sensor nodes of very small sizes having sufficient computing power at affordable costs. These developments have made the use of these devices very popular. As such applications of WSN where sensor nodes are the basic constituent have invaded every walk of life. Today we can find WSN applications into the areas such as healthcare, traffic monitoring, remote control of appliances in buildings, surveillance in difficult areas etc. As a result of sensing the area in which these networks operate a lot of data is produced. Millions of bits of data in the form of continuous data streams become available for manipulation to extract desired information.

A lot of literature related to evolution of sensor networks from limited defense or specialized applications to every walk of life, their limitations and distinctions from the traditional networks as a result of these limitations were reviewed. Quest to utilize the limited WSN resources optimally has driven the researchers world over to undertake numerous research projects to enhance resource efficiency. The research has resulted in development of a number of variants of network protocols for network operations, techniques and means of data extraction. These means and protocols are kept application specific as the main target is to achieve maximum for the metrics of significance for that particular application only.

For reliable and economic extraction of information various database management systems/ techniques have been developed. The objective is to manage the ever growing data so as to facilitate information extraction with minimum drain on resources. These processes/ techniques may be categorized into three broad categories [23-25]: *agent-based*, *query-based* and *macro-programming based*.

## INFORMATION EXTRACTION TECHNIQUES

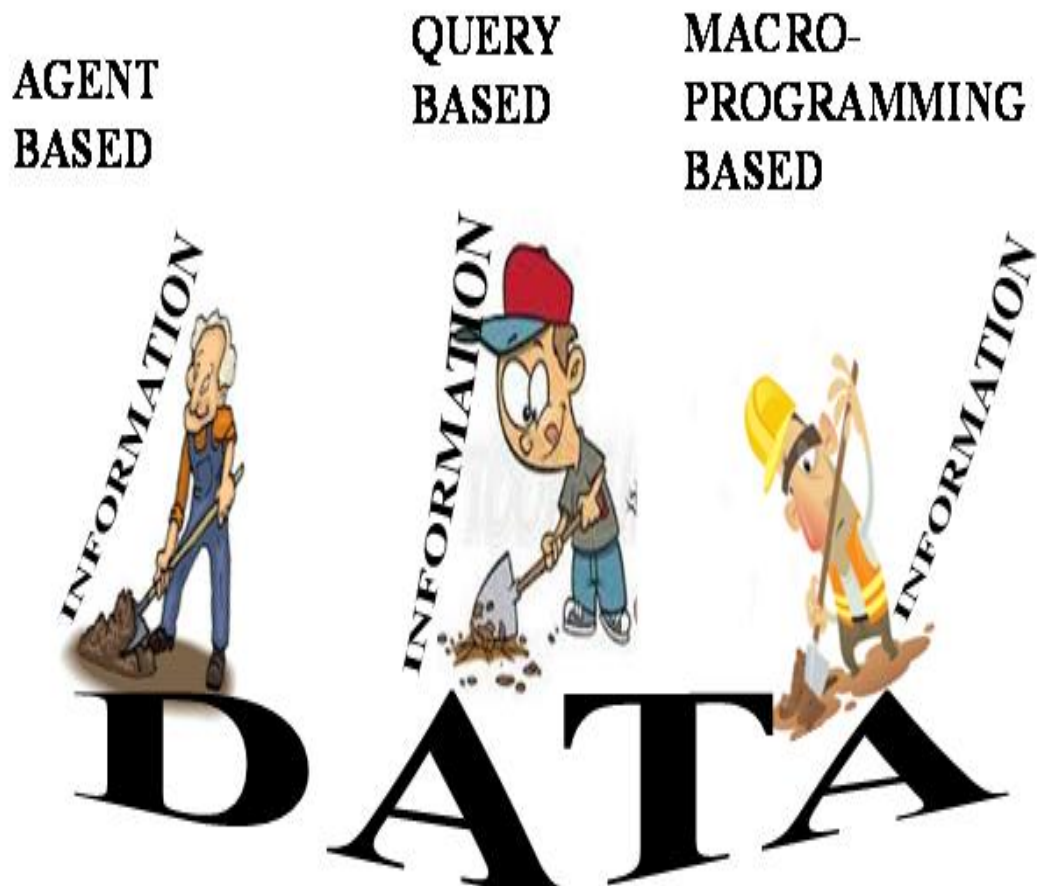


Figure 2.1: Various Information extraction techniques

Literature has been reviewed and a brief of the finding is-WSN may be abstracted as a distributed database of sensor data. Therefore, query based approach is very easy and popular. Research has been focused on devising efficient query processing schemes or techniques.

Till early 80s query processing techniques were mostly available for wired systems only. Researchers wanted to broaden the applicability of these techniques to wireless networks also, so that the use of wireless networks could be made popular. WSN is built up of sensor nodes deployed over a large geographical area. Each node in a WSN acts as a source of input from where the sensed data makes its way into the WSN for reaching the user. Each node within a WSN is having sensed data and this can be considered as a mini database which is *distributed* in nature, making WSN a “*distributed database*”. To this distributed data, queries are to be applied to extract the desired information. Query processing techniques may vary so as to meet the QoS and performance requirements. A number of evaluation metrics are employed to compare these techniques and best possible is chosen.

Some of the metrics used for evaluation of performance are *timeliness*, *accuracy*, *reliability* and *optimum usage of resources* etc. These metrics are not independent of each other. To achieve better score in one parameter, we may have to sacrifice some other parameter. Therefore desired QoS is the main objective for which a particular metric may have greater significance over others.

As costs of wireless networks were getting down, the versatility of nodes made WSN suitable to a number of applications. Moreover the WSN could be set up easily without any previous infrastructure. Initially wireless networks were used for some specialized services such as defense etc. as data retrieval methods from these networks were not developed and costs of hardware were also prohibitive. WSN have caught the fancy of many research and development scholars recently due to increasing affordability and versatility of these networks. Applications of WSN have grown, however the inherent limitations such as limited power, low processing capabilities and available bandwidth hamper the growth of WSN applications. There is a whole lot of research going on in WSN field with the main emphasis on energy efficiency.

## 2.1 Agent based approach

Agents are autonomous software which work independently for achieving any designated task without the intervention of user. They are equipped with capabilities of decision making, self-replication and handling unpredictable environmental so [26,27] which make them suitable for WSN environment. The agent based systems may have a group of agents, which team up for executing a particular task for information extraction which is a feature of WSN in most of the applications. Agents may be static or mobile.

H. Qi et al.[28] proposed Distributed Sensor Network (DSN) architecture having mobile agents. Mobile agents transmit the computation code to the sensors having raw data. This saves energy used in sending raw data to base station. Moreover the system remains unaffected by intermittent fluctuations in data or connectivity. Robert Wesson et al.[29] in their work have shown 90% savings in data transfer by using Mobile agent DSN. Similar model of In-Network processing incorporating collaborative signal and information processing discussed by Hairong Qi et al. in [30] again has shown a large amount of saving.

Bhupinder Singh et al [31] have introduced mobile agents which are guided by ant's colony optimization for self-adaptation to suit dynamic nature of wireless networks. The mobile agents are able to discover required rate from heterogeneous, dynamic networks.

Work by C. Min et al. [32] has shown that in a multi hop network having flat architecture, mobile agents may be designed to use context aware In-network processing or data aggregation to reduce communication load. In this approach data redundancy can be eliminated at node level or task level.

An agent based framework has been proposed by L. Szumel et al. [33] which can be helpful in a wide variety of applications. In this scheme agents are mobile and can visit one or more nodes. Agent uses the resources of the node which it has moved. Agent based approaches are delivering encouraging results however the distributed databases of WSN pose a challenging task in designing the mobile agents because the resource constraints and unforeseeable environmental conditions of WSN are to be taken care of.

Moreover testing and debugging agent based system is not easy. Some of the major frameworks using agent based approach are- Mate [34], Agilla[35], and Impala [36].

## **2.2 Macro programming based approach**

The second approach is to develop macro program for WSN. In macro programming WSN is treated as a whole by the programmer, individual node level programming is not required. It is based on a high level language, which incorporate global abstractions. Node level data is abstracted by modeling and is approached by the macro programs. In this the programmer is relieved from the in depth study of the system's resources, topology etc. On the contrary the programmer is required to possess sufficient amount of knowledge and experience so that a macro program engulfing large heterogeneous networks of wireless nodes can be designed.

Whitehouse et al. [37] proposed node level abstraction called hood in which neighborhood nodes are selected on the basis of set criteria and the set of variables which are to be shared. Range of hood may vary according to application or parameter required. A hood can be one hop for an attribute and two or more hops for other attributes.

Although macro programming is supposed to save the programmer from the programming at the node level, however certain macro programming based systems cannot fully abstract the details up to the node level. In such cases node level knowledge becomes necessary. Moreover, learning of macro programming is time consuming and requires a lot of expertise.

Some of the major systems using macro programming based approach are- Kairos[38], Knowledge-Representation for Sentient Computing[39].

## **2.3 Query based approach**

This is the most widely used approach for data extraction. It is based on querying the database for information retrieval.

All the three approaches are well suited for diverse WSN applications. Undoubtedly, agent based approach and macro programming approach have an upper hand in terms of flexibility but implementing them in WSN is not an easy task in comparison to query based approach as the user feels a great degree of comfort while working with queries. Therefore because of popularity and ease of use query based approaches have been reviewed.

Through queries a user can get timely and reliable information economically from very large database systems running in diversified environments. Query instructs a DBMS to retrieve or update data. Complete processing of a query generally involves three steps: Parsing or translation followed by Optimization and finally Evaluation. The first step involves conversion of query into a form understandable by the query processing engine. In the next stage the query processor converts the internal data structure into equivalent but more efficient representations. This internal representation by query processor may be done on the basis of mathematical or cost models, heuristics, semantics or selection algorithms etc. Choosing these and applying them for an efficient result is the work of Query Optimization Engine. It develops plans which are evaluated in the final phase. Finally, the evaluation phase selects the best plan out of the various plans generated by the optimization engine keeping in view the application specific metrics such as latency, energy efficiency, security etc.

Some of the major frameworks using query based approach are- COUGAR [40], Sensor Information Networking Architecture (SINA)[41], Data Service Middleware (DSWare)[42], Framework in Java for Operators on Remote Data Streams (Fjords)[43] and TinyDB[44].

## **2.4 DBMS Query Processing vs. WSN Query Processing**

Query processing in traditional databases and in databases related to WSN differs to a great extent because of the inherent characteristics of these networks and nodes. Query

processing plans in traditional databases are mostly unsuitable for WSN. In traditional databases, the best query plan chosen by query optimizer is the one which needs minimum number of disk accesses. However, in WSN the best query plan chosen by the optimizer is the one which has minimal estimated energy cost. In addition to this the design limitations of the sensors constituting the wireless network force architecture and protocol stacks of WSN to be different from traditional networks. The nodes comprising a WSN are more prone to failures because of harsh environment and low costs. Therefore, nodes have restricted memory size, use vast amount of energy for data transmissions which is in the continuous data streams.

Table 2.2: DBMS Query Processing vs. WSN Query Processing.

Type of Database	DBMS(Traditional) Database	WSN Database
Features		
Processing Level	Centralized	Centralized; Distributed
Data Type	Static	Real-time continuous data stream
Available Memory	Available without any constraints	Restricted
Computational abilities	Unlimited	Restricted
Available Energy	Unrestricted	Limited

Therefore, in contrast to the traditional query processing in DBMS, the input received from a sensor of a WSN is in the form of continuous data streams in an unordered and unreliable form (due to various environmental factors). Therefore, in order to ensure efficient query processing in WSN, resource management like- memory, energy, bandwidth etc. should be taken care of. Query processing techniques in sensor networks may be classified as:

- a) *In-network* or
- b) *Base station* or as distributed processing or warehousing.

P. Bonnet et al. [45] states that Distributed processing or In-network processing reduces transmission costs as raw data is processed at node level and data to be transmitted reduces considerably.

Warehousing approach involves data transmission (for processing), to a place centrally located equipped with abundant energy. Main disadvantage associated with this approach is the vast amount of energy requirements for data transmission from node to the centrally located area. Moreover, even when information is to be extracted from a small portion of the data, the warehousing approach follows the order of transmitting the whole data to the central location for processing and analysis. Main advantage is that there is centralized control which saves retransmission costs which is very much likely in WSN environment.

## 2.5 Energy Efficient Approaches for WSN

Easy availability of WSN which are suitable for almost any thinkable application has made energy efficiency or conservation a crucial issue. Various approaches keeping in view the priority of desired metrics (defined earlier) have been devised. These approaches may be classified into three categories namely *mobility based*, *database optimization* and *duty cycling*.

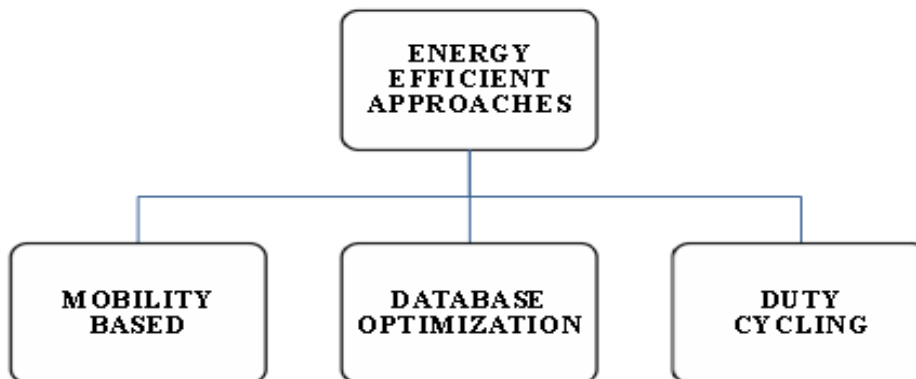


Figure 2.3: Various Energy efficient approaches in a WSN

### **2.5.1 Mobility based**

The data sink is made to travel to node in a network for data collection reducing the energy consumption. The mobility based approach used for data collection is divided into two categories i.e.

- i) *using a single mobile sink* and
- ii) *using multiple mobile sinks.*

In single sink approach the data packets are collected by the sink itself resulting in huge benefits. However, the failure of single mobile node may lead to loss of precious data as data is generally real time data. To overcome this problem, multiple mobile sinks are used for data collection. The mobile sinks may collect the data from the nodes by moving in a random, predictable [46] and controlled mobility fashion [47]. In this way risk of data loss in case of single failure is eliminated. However, this requires more detailing or even some redundancy may also crop up.

### **2.5.2 Duty Cycling**

In this mode, sleep and wake cycle called *duty cycle* of a node is adjusted according to the application thereby reducing wastage of energy in a node when it is idle. Sensors consume a large amount of energy in idle mode also. Therefore, the nodes which are used for sensing, reception and transmission of data are switched off when idle. This saves wastage of energy while idle. However, downside of this approach is that a node may fail to detect some event when it is sleep. Moreover, switching off of radio signals intermittently may lead to connectivity problems. Hence a tradeoff is to be maintained. By operating at a low duty cycle i.e. cycle of operation of a node which operates intermittently helps in energy conservation. It is most appropriate in conditions where sensors usually remain idle. A trade-off has to be established between the sensing and data transmission in an intermittent manner in WSN, as nodes go into sleep and wake mode. The intermittent condition of the node sometimes fails to detect an event. Also switching off the radio signals may lead to connectivity problem between the nodes. As a result the data is unable to find the path for its transmission.

Sleep awake scheduling is an effective way to reduce energy consumption of a node but introduces delay as each node has to wait for the receiver to awake. In delay sensitive applications this delay is unacceptable. In case of synchronized sleep/awake scheduling [48-50] sensor nodes exchange synchronizing information with neighboring nodes, this incur additional communication load. Therefore asynchronous sleep/awake protocols [51,52] are more efficient. However to reduce the delays Anycast has been proposed in [53-57]. In this the sending node wakes up a group of neighboring nodes and picks up the node which has woken up first. This reduces the waiting time and hence delays. However the set of nodes to be awakened is to be chosen carefully as it may affect routing path also.

To maintain reliability in case of node or communication failure, redundant nodes or paths are set up in a WSN. Node redundancy can be used to advantage and only a subset of nodes is made active and put to use [58-60]. These active nodes cater to the need of monitoring whole area and maintain connectivity. By shifting different subset to active/ sleep state energy can be conserved. However, if node redundancy is lost due to any fault this mechanism is not useful [61, 62].

### 2.5.3 Database Optimization

Optimization of database consists of optimization at *data acquisition* and *data processing stage*.

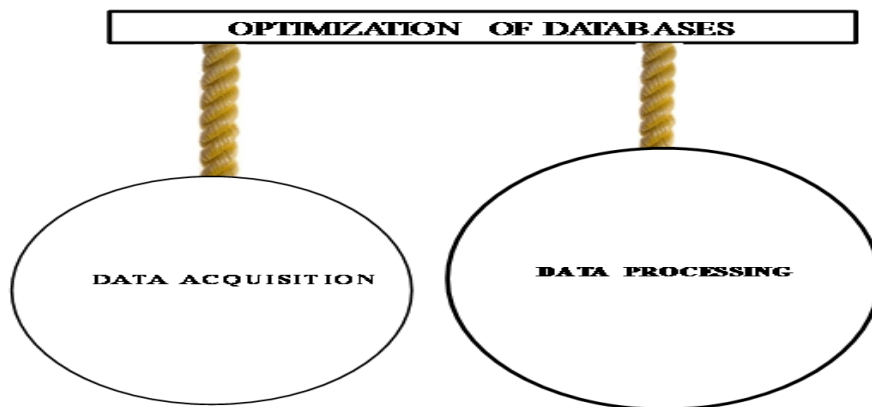


Figure 2.4: Various Database Optimization approaches in a WSN

### 2.5.3.1 Data Acquisition

Data acquisition is the process of measuring or sensing the physical parameter or event of interest converting this data into digital numeric values which can be operated upon with the help of computer. Physical parameter that is to be measured is initially sampled with the help of sensors. Measured signals are analog in nature. These signals are converted into digital form so as to make operable by the microcontroller. Elements of data acquisition system are: transducers, signal conditioners and ADCs. Data acquisition may be summed as acquisition of signals from the environment, subjecting them to digitization and finally saving as digital data.

Traditional Data acquisition approaches can be categorized as: *Data logging* and *Decentralized data aggregation*.

In the first case the data is acquired at various nodes and then each node transmits the data to the BS. It is a time consuming and an inefficient approach but limited bandwidth is used efficiently in this way. However, in decentralized data aggregation the acquired data is processed at various nodes and aggregated before its transmission to the sinknode as has been described by Rice et al. in [63]. Processing of data at node level decrease the size of the data to be transmitted [64] there by reducing energy consumption. According to C. Alippi et al. [65] energy efficient data acquisition can be achieved through techniques like *duty cycling* and *adaptive sensing*. In Duty cycling sleep and wake cycle of sensor node is adjusted. Adaptive sensing/ data acquisition can be done with the help of any one of the below mentioned techniques i.e. *hierarchical sensing*, *adaptive sampling* and *model-driven sampling*.

*Hierarchical Sampling:* In this approach the sensed data moves through a path from energy rich nodes (with low energy consumption) to the nodes possessing lesser amount of energy lying in close proximity to the BS. Low energy consuming nodes are given priority in hierarchy.

*Adaptive Sampling:* It is based on the relationship existing between the sensed data and the energy possessed by the node. Temporal correlation exists when there is not much change in the sensed data spanning over a period of time. Spatial correlation exists when the sensors lying in close proximity sense the data with no distinction/ dissimilarity. Hybrid of both temporal and spatial correlation may be used for getting the samples, keeping in view the energy left within a node.

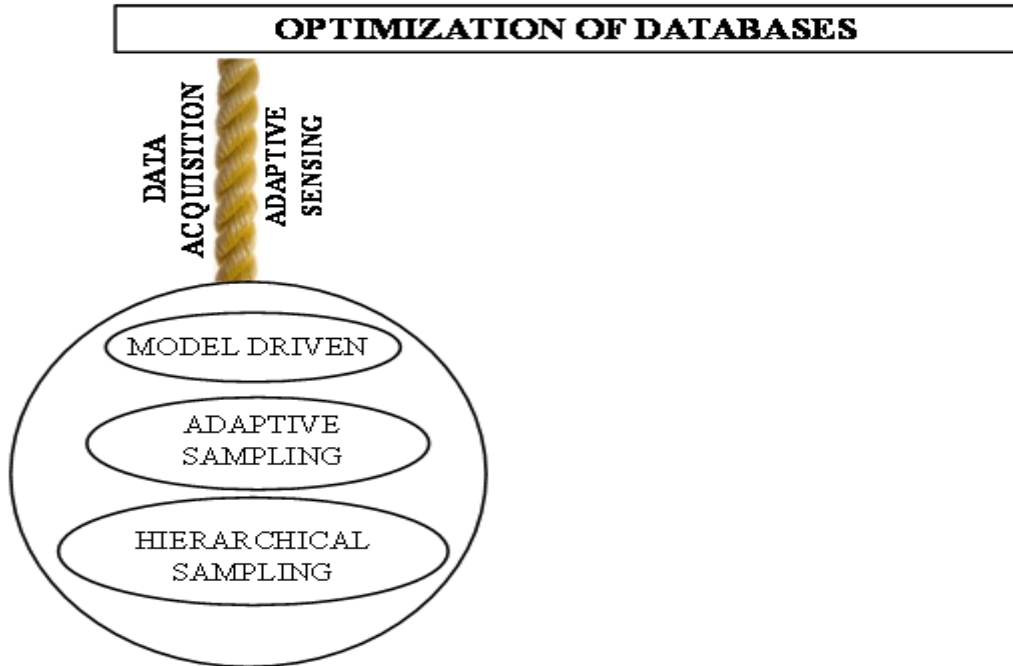


Figure 2.5: Various Adaptive Sensing/ Data Acquisition approaches in a WSN

*Model driven:* In this approach a model is constructed with the help of sensed data over a period of time. This model is used to predict data. The approach is energy efficient as continuous data transmissions may be avoided (as data is predicted). Only changes beyond a certain band from the predicted values are transmitted. Model based approaches require regular model updating to comply with the changing physical conditions of the area under study to maintain reasonable authenticity. In addition to this only approximate values are available because available data is based on model driven prediction. Sangeeta Rani et al. in [66] have used data series clustering for understanding inner trends of time series dependence of data which will help in predicting the future values of data.

Sally Ida McClean et al. in [67] have introduced new term activity recognition which uses low level sensor data to construct higher level activities behavior of network. This helps in creation of models for activity prediction.

### 2.5.3.2 Data Processing

Different data processing strategies are required for different applications. Data processing needs may change with passage of time foreven the same applications. As user interests may change and user may require different data over a passage of time or user may demand data on event occurrence only. In such cases SungboSeo[68] has proposed multivariate stream data classification where signal changes are differentiated by assigning different symbols and then classification algorithms are applied to classify the data. Various approaches of data processing are as follows:

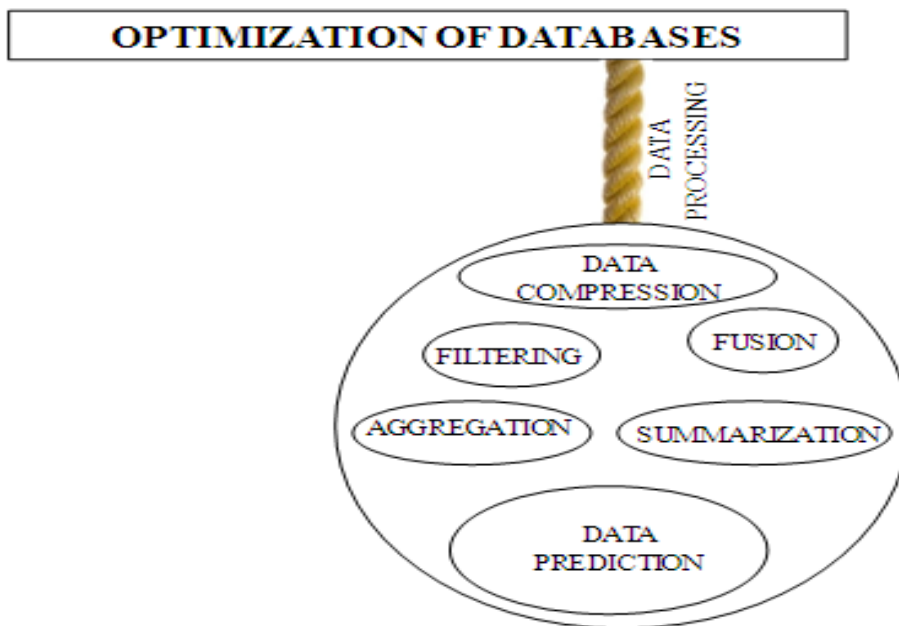


Figure 2.6: Various Energy Efficient Data Processing approaches in a WSN

- a. *Data compression*: It is one of the basic measures adopted to optimize usage of limited resources. Data compression leads to reduction in size of the data to be transmitted, which require- less bandwidth and less power. Though power is

consumed in executing the compression algorithm also, however, there is a net power gain as the cost of data compression is less than the cost of data transmission. Compression algorithms considered for the sensed data in a dense sensor network follow a *distributed approach* here as in sparse sensor network *local approach* is implemented. The advantages of using compression technique are that power optimization leads to increase in WSN's performance; reduction in cost, space & data transmission time; etc.

Keeping in view the desired QoS, various techniques and algorithms are in use for data compression. Huffman code [69] is a lossless compression technique. Video/ image compression uses techniques such as JPEG 2000 [70] which involve some loss.

Techniques of data compression and associated data loss:

- i. String based compression: provides lossless compression.
- ii. Image based compression: Some minor loss is present.
- iii. Distributed source coding: It uses Seepian Wolf Theorem which states that two or more co related data streams can be encoded independently and decoded jointly at the receiver without any loss.

The *Lempel–Ziv–Welch* (LZW) algorithm [71] constructs a dictionary dynamically to encode new strings based on previously encountered strings. LZW is computationally simple and has no transmission overheads. A special algorithm for WSN SLZW called sensor LZW has been proposed in [72]. Other variant of LZW such as SLZW-MC is also available.

In image based compression, WSN is organized into a hierarchical architecture and sensor data is considered as an image. Wavelet transformation is performed to extract spatial and temporal correlation.

- b. *Filtering*: Due to sensing, a node generates data continuously which is to be handled efficiently as per the requirement of the application. Besides the genuine useable data, a node may get afflicted with false data injection from rogue nodes. This rogue data is a major threat to network security and energy resources.

Effective filtering techniques must be in place to identify and separate the rogue data and rogue nodes also. Filtering technique “Bandwidth Efficient Cooperative Authentication” (BECAN) [73] is used for early detection and filtering of maximum possible false data. Source authentication to identify rogue nodes injecting false data is also employed. This technique has high computational and communicational overheads of its own which is a drain on energy. Schemes with low overheads are also available, tradeoff between energy depletion due to overheads and desired QoS is to be decided by the user.

- c. *Data Fusion:* WSN are deployed to monitor and measure physical conditions, which may be very harsh and dynamic. Sometimes strong variations in ambient conditions affect the measurement of sensors resulting in acquisition of imprecise or unreliable data. Data fusion techniques are employed to overcome such problems. In this, data from multiple sources is combined and analyzed for its validation. It helps in getting more accurate data as compared to data given by single source. Data fusion can be applied on centralized or distributed basis. In WSN, distributed fusion means fusion of data of a node and its neighbors. Data fusion is helpful in overcoming effects of sensor failure, technological limitations and spatial and temporal coverage problems through cooperation, redundancy and complementarities. Sensor nodes deployed in a network may cover a partial portion of the network. Fusion of data of nodes cooperating with each other results in complete view of the network. Redundancy helps in mitigating the effect of sensor failure and complementarities is used to make different sensors observe different parameters and fusion is used to draw out inferences which may not have been possible with single measurement. As such researchers [74,75] have defined data fusion as “multilevel, multifaceted process handling the automatic detection, association, correlation, estimation, and combination of data and information from several sources”. Fusion of data decreases probability of errors and enhance reliability of data. Increased confidence in reliability and accuracy of the acquired data helps in better understandability of WSN application resulting in execution of response quickly.

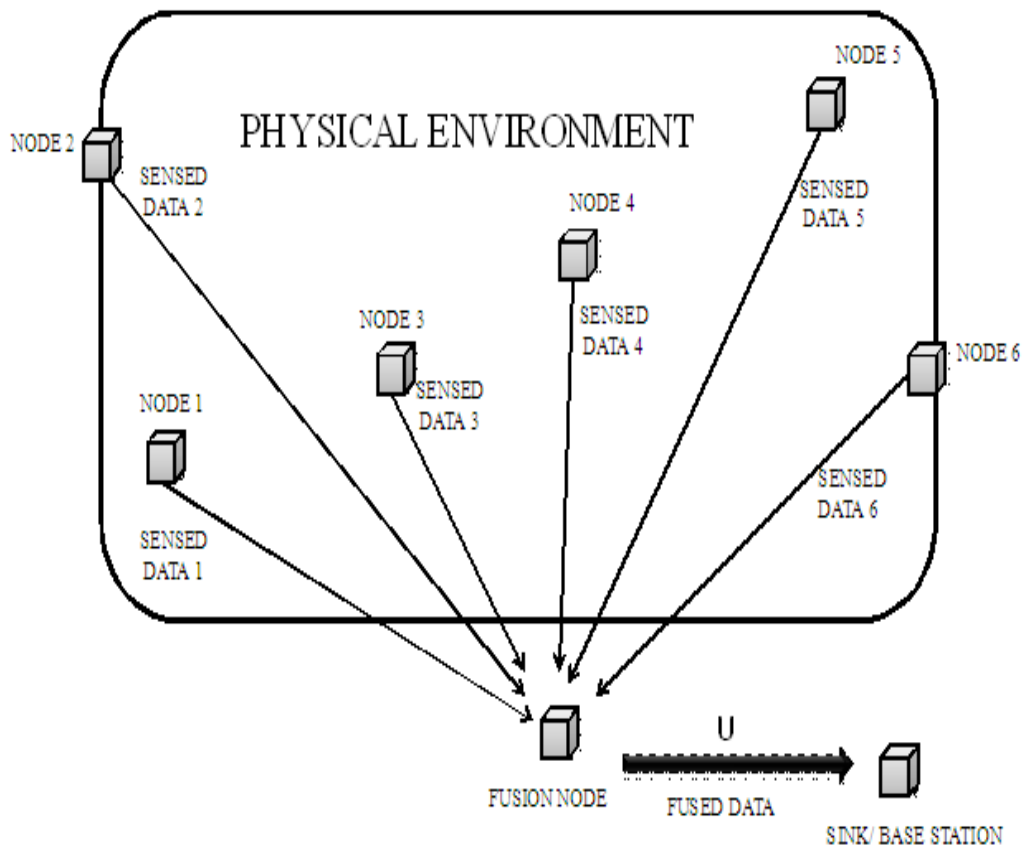


Figure 2.7: Data Fusion- Collecting results from multiple nodes

On continuous data stream, Data Fusion is to be implemented at regular intervals, Otherwise it may result into data explosion. Various models based on- input and output obtained [74], in context to software engineering [76] and random sets [77] have been proposed for data fusion. Data fusion is a tough job to handle as the available techniques vary in range and have many other limitations also.

- d. *Data aggregation:* The terms Data Fusion and Data Aggregation seem to be the same therefore; there is lot of confusion about the interchangeability of these terms. No doubt that there are similarities between the two but the protocols designed for them differ in terms of the number of nodes deployed (density) and the frequency with which the queries are injected. Data fusion is applied on a WSN covering a large area using compressive functions where as data

aggregation is applied on a WSN covering a small area using functions like count, sum, average etc.

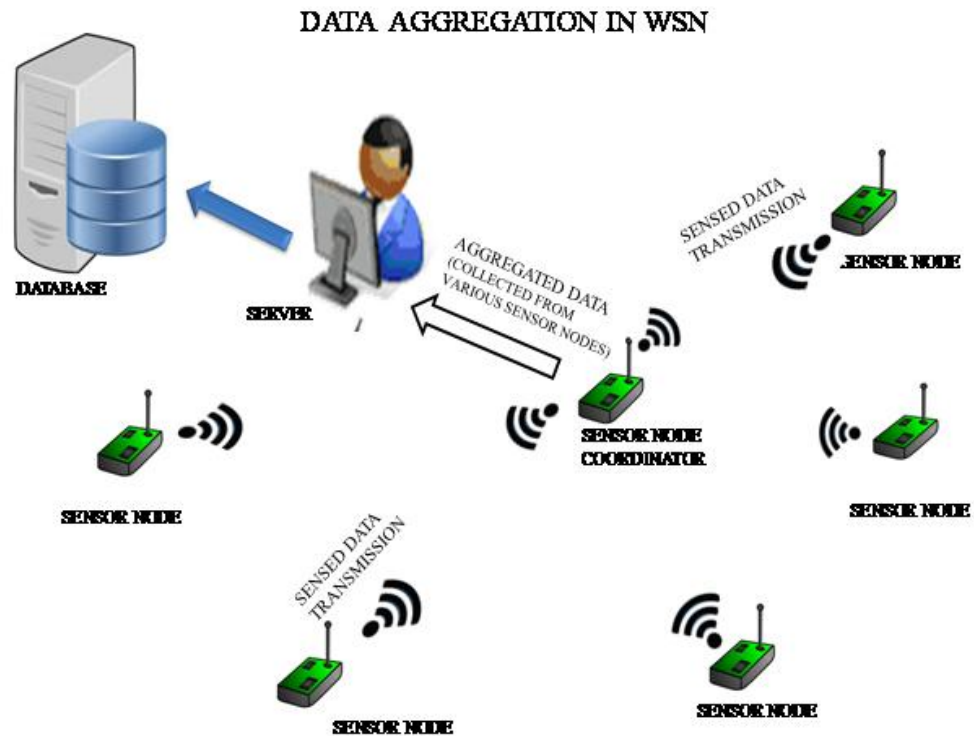


Figure 2.8: Data aggregation performed by Sensor Node Coordinator

Data aggregation aims at increasing the lifetime of the network by reducing data transmissions which are redundant. This technique aggregates the similar data sensed by various sensor nodes while passing through various nodes before reaching the base station. The data may be aggregated with the help of various aggregation schemes like Low Energy Adaptive Clustering Hierarchy (LEACH) [78], Tiny Aggregation (TAG) [79] etc.

The aggregation may be performed either by

- i) *Reducing the data size* or
- ii) *Without data size reduction* that is to be transmitted.

The former method involves combining and compressing the data sensed at a particular node which receives the data from other neighboring nodes before allowing the same to

make its way to the BS/ sink. In the second method, the data that is received from other nodes is merged at a particular node which reduces communication overheads. No processing takes place on this merged data which is then forwarded to the base station. Data aggregation can also be achieved through specific routing schemes such as:

- i) *Address based* and
- ii) *data based.*

In the address based routing scheme [80] the query is sent to the particular address and the node of specified address in the query transmits the result to the BS. In contrast to the address based routing, in data centric routing, the query is broadcast to all the nodes within the WSN which satisfy the given condition as mentioned in the injected query. All the nodes which have that similar condition combine their data and aggregated data makes its way to the base station.

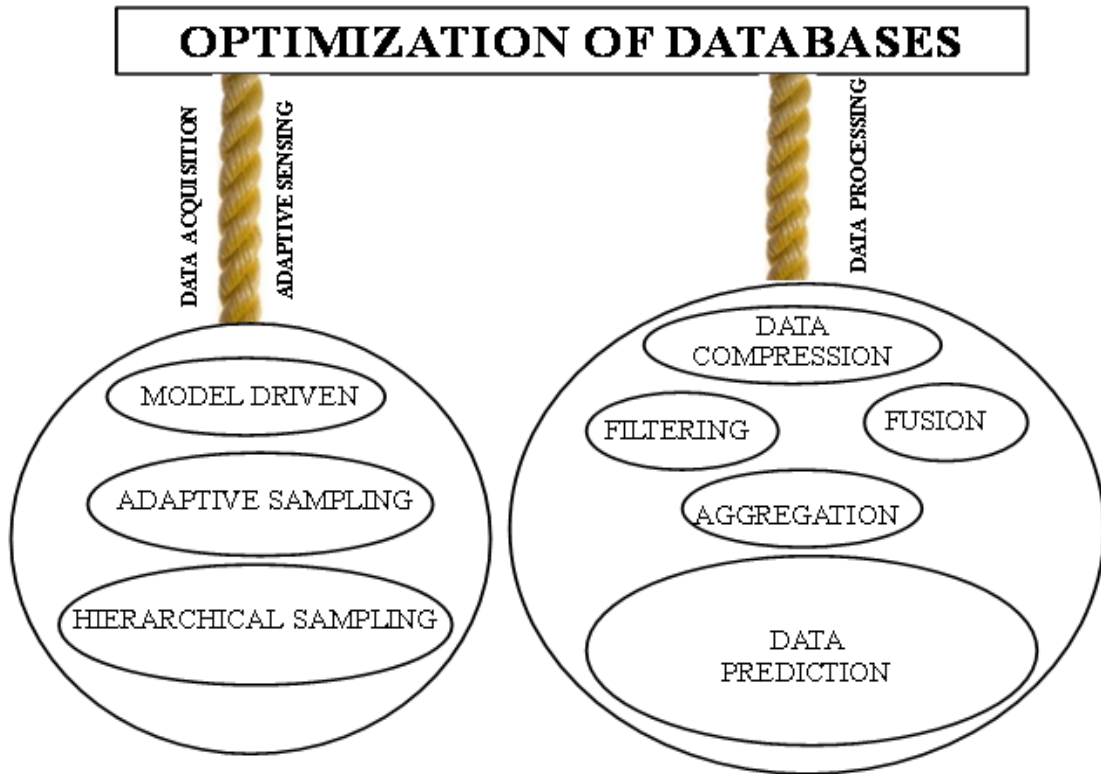


Figure 2.9: Various Energy Efficient Database Optimization approaches in a WSN at a glance

=

- e. *Data prediction*: Prediction based data reduction/ optimization is very helpful in WSN as environmental data sensed periodically has a great temporal redundancy. Prediction based approaches exploit temporal correlation of the sensed data and prepare model for data prediction. This saves the network from communicating redundant data or data within the acceptable accuracy. Due to changing ambient conditions the model needs to be updated regularly. This model updating involves its own overheads which need to be studied to work out the tradeoff point. On the basis of above study we may conclude that optimization of database can be summarized as in the Figure 2.6.

## **2.6 Types of Query processing Frameworks**

WSN can aptly be called a distributed database because of the availability of the data in the nodes distributed over the network. Though inherent computing capabilities of these nodes are very limited, however effective gains can be made through use of these limited processing capabilities also. Processing of query or data in these nodes can prove to be a major energy saver.

Query processing at node level is *Distributed data processing approach*. Intanagonwivat et al. [81] conceived the idea of the term “In network processing” in WSN.

The credit of being the first generation systems employed for query processing in WSN goes to TinyDb[44] and COUGAR [40] which use SQL.

Various frameworks have been proposed for various types of queries and various processes and the techniques employed in processing those queries.

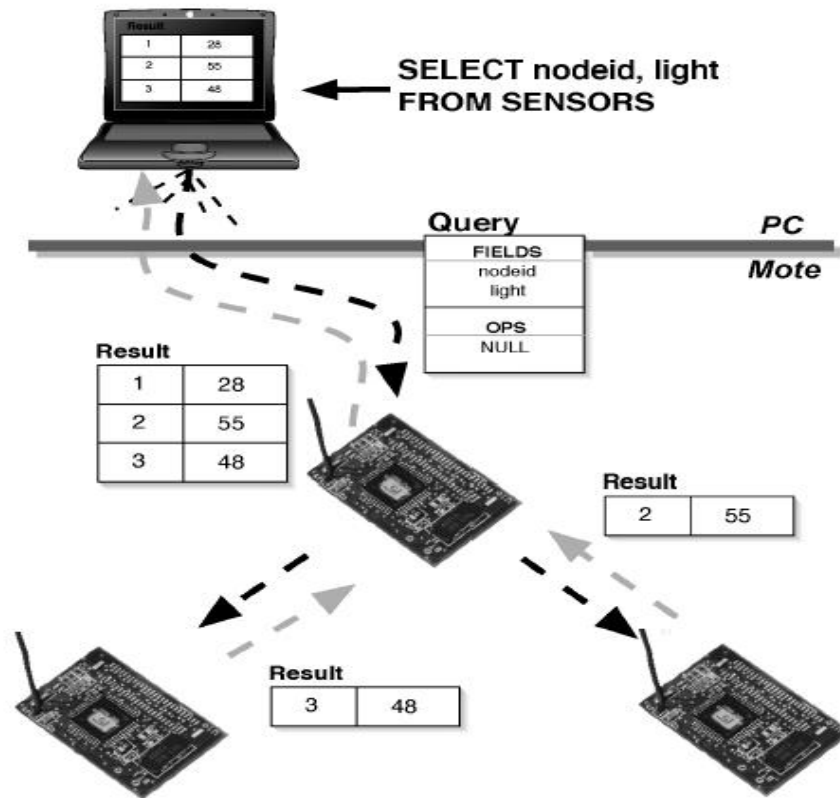


Figure 2.10: TinyDb approach: Query Processing Model [44]

TinyDb[44] and COUGAR [40] use distributed approach to query processing. SQL based query is injected into the network. Query is routed through a spanning tree with sink as root. Results are sent back to the sink through the tree. Data aggregation on data is carried out at nodes enroute, in this way, irrelevant data is discarded. In COUGAR [40] query proxy layer is placed on each node. Query optimizer placed on gateway node, distributes query processing plans to the network nodes. Query processing plans giving detailed plan of data flow and information processing at each node are constructed by COUGAR on the basis of metadata catalog and query specifications. It was found that TinyDb overweighed COUGAR in context to energy savings. In SINA [41] hierarchical clustering of sensors on the basis of available power and proximity is done to increase efficiency. Cluster heads process information by filtering, fusion and aggregation. For accessing the data it employs Sensor Query and Tasking Language (SCTL). Both event

monitoring and data accessing according to the query injected are its key features. It adopts methods like sampling operation to make it resource efficient.

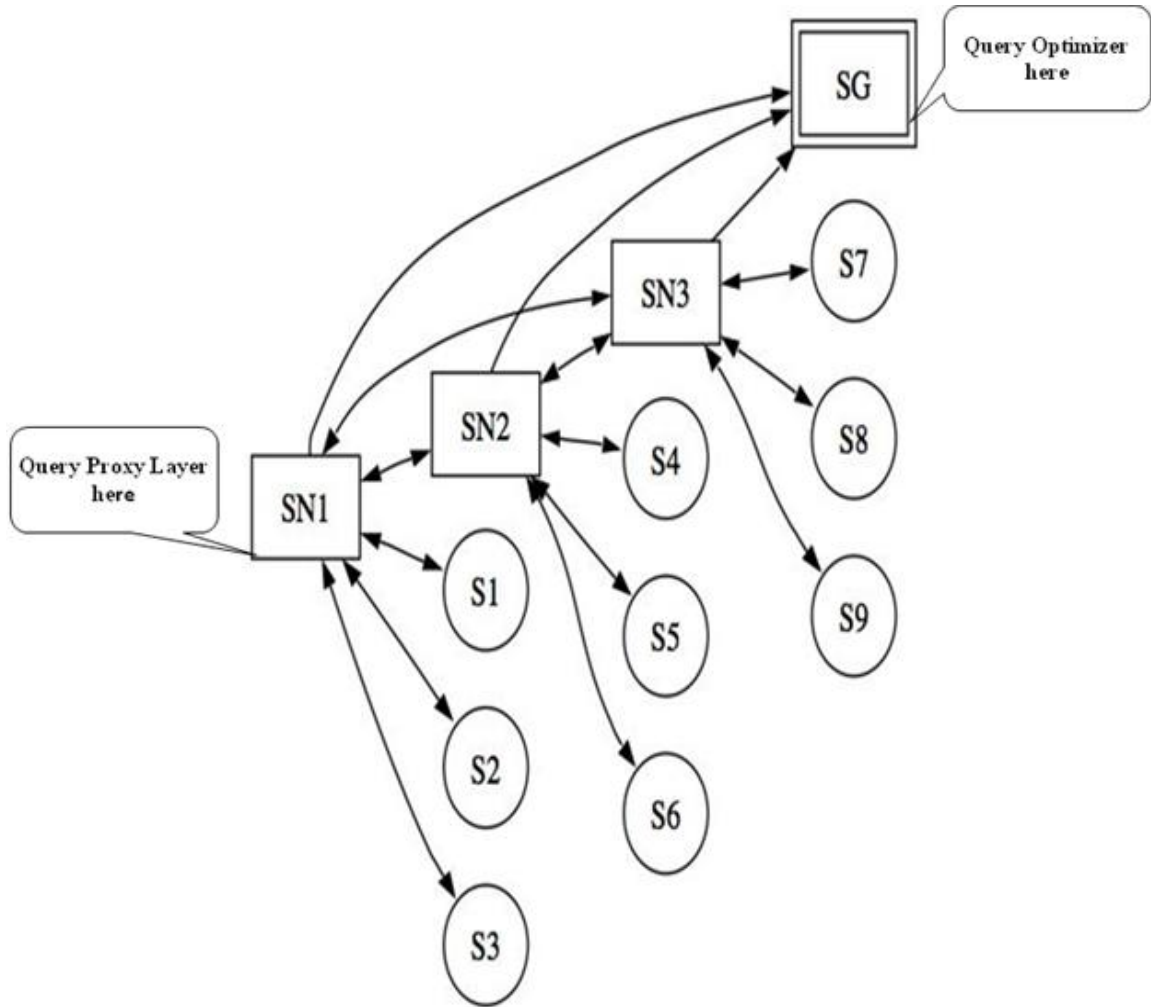


Figure 2.11: The COUGAR approach [40]

These systems provide high level interfaces and inner workings of the network are abstracted. However, the range of queries that can be issued in these systems is very limited. Queries cannot express spatio-temporal characteristics also. These limitations can be overcome by finding new low level operators which means application specific functionality has to be introduced which comprises the feature of inner network abstraction. .

In subsequent work it was found that the focus of concern as stated by Gehrke et al. [43] was multi-query optimization, choosing nodes which stored the concerned data and heterogeneous networks. SINA [41] too follows distributed database query interface. However, SINA failed to address the problems of spatial/ temporal queries. P. Bonnet et al. in [82] considered WSN to be a database. This database was equipped with warehousing and distributed query processing approaches. The proposed method was not apt for processing continuous data stream which is basic characteristic of data generated in WSN.

Active Query Forwarding (ACQUIRE) [83] too is in the category of distributed data processing. Here a query is injected into WSN via a node. Nodes lying in close proximity to the node that injected the query transmit updated information to that node. The node further picks up another node for broadcasting that query in order to get the desired result. The system can be used for nested queries also. However, spatial or temporal characteristics of a query are not addressed.

A two-phase algorithm for query routing and result assimilation & processing of query formed the basis of the framework as proposed by A. Coman et al. [84]. In another In-Network Querying framework for WSN proposed by M. Demirbas et al. [85], the difficulties encountered due to the resource constraints while executing the queries, were addressed by making use of Distributed Quad-Tree (DQT).

Data Service middleware (DSWare) as stated by L. Shuoqi et al. [42] is a data centric approach for real-time applications. It is a middleware approach which can handle real time events and at the same time it can provide abstractions which are available in case of databases. The system is not suitable for complex aggregate queries.

Framework in Java for Operation on Remote data Streams (Fjords) proposed by Madden et al. [43] proposes an architecture which supports processing of multiple queries over continuous data streams generated by sensor nodes.

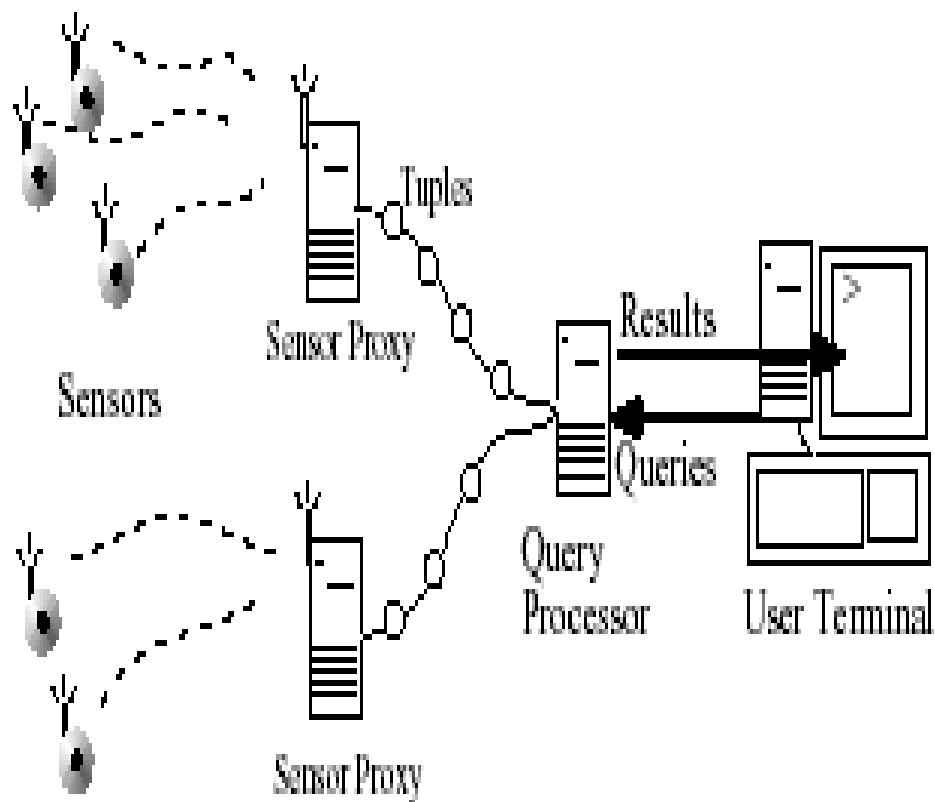


Figure 2.12: Fjords architecture [43]

Fjord operators known as sensor proxies serve as mediator between query processing plan and sensor. Fjord operator waits for data from the sensor and pushes it to the query processor. Fjord operator acts as non-blocking operator as it receives data from the sensor as and when it comes and pushes it to the query processor. These operators receive and push data therefore operators have time to assess other parameters of the network such as power, communication bandwidth etc. in this way it addresses issue of energy efficiency to some extent which is ignored in COUGAR. However, it is not able to handle complex queries or aggregate type queries.

Y.Chen et al. [86]proposed “an energy-efficient framework for multi-rate query” where only a single data source is available to serve information needs of queries emanated

from various base stations requiring data at different frequencies. The framework is one of its own kinds where an integrated data stream is communicated over high frequencies from which various data streams of desired frequencies can be reconstructed. This requires rate conversion mechanism and elaborate path sharing routing techniques.

Wringer [87] gave a framework that employs tools for finding the bugs and fixing them was developed by A.Tavakoli in 2007. RUASN (A Robust User Authentication framework for Wireless Sensor Networks) by Pardeep et al. [88] was a framework that dealt with the security aspects. H.Jiang et al. [89] proposed a framework based on clustering along with prediction scheme where the network is divided into various clusters headed by a cluster head which is responsible for data collection from their respective clusters. The authors claim their work to be an initiative in the field of prediction scheme for clustering-based data collection in WSN.

“Hybrid in-network Query processing Framework for WSNs” as proposed by S. Pervin et al. [90] combined both approximate and aggregate query processing techniques. While doing this it was ensured that their constraints were overcome. A framework proposed by AmroQandour[91] is developed to render help in monitoring the physical conditions like the gas leakage, quality of air etc. To enhance the affinity between the physical world and the WSNs, the framework was equipped with OverThe Air Programming (OTAP), the connections within the network, sensors’ governability and alarm signaling. LambrosSarakis et al. [92] worked in an emerging field of VirtualSensorNetworks (VSNs) in the area of WSNs. Efforts were made in realizing the scalability, flexibility, energy-efficiency with the aim of decrease in cost, deployment, complexity etc.

MAC designed for traditional network laid less importance on energy conservation. So many new MAC protocols have been designed which have WSN specific features. MACAW [72] and IEEE 802.11 [93] eliminate wastage in collisions. Further energy loss has been checked by improvement in PAMAS[94] where unintended destinations are not sent the packets. It has been seen that idle power consumption is also a significant portion so MAC protocols for sensor networks control sleep awake cycle also.

Jose E. Bessa Maia et al. [95] in their work proposed a framework for processing complex queries where the framework proved to be energy efficient by decreasing the

amount of data injected into the network. EvAnT framework proposed by M.Woehrle et al. [96] is used for analyzing the events in the network. The framework supports heterogeneity in terms of devices and test beds. The framework works towards enhancing the life-time of the network. Dynamic Conflict-free Query Scheduling (DCQS) is another framework proposed by O.Chipra et al. [97]. The framework developed claims to provide conflict-free transmission. Its dynamic nature adjusts itself according to the changes in workload without involving much change in the transmission schedule.

After extensive study of literature regarding evolution, applications and efficient operation of WSN as discussed above, it has been inferred that:

- I. WSN have a great potential to serve the human need of anytime and any information.
- II. WSN can be scaled up to cover large geographical extent and can serve multiple applications of multiple users.
- III. Query based approach is the most popular and easy approach of data extraction in WSN.
- IV. Resource constraints of these networks are forcing research community to develop and devise means which use the resources judiciously.

In WSN, efficiency of one parameter may come at the expense of some other parameter. Therefore, most of the research activity has been application specific so that significant parameter for that application is available efficiently.

However, WSN are able to serve multiple applications and many of these applications may have some correlation which is not exploited by the application specific techniques or tools. It results in processing and transmission of similar data multiple times in response to various applications. Therefore, in data intensive WSN it is imperative that same generalized approach be used which treats whole data and network and exploit correlation and redundancy of data so that processing and transmission of data is not repeated for similar queries.

The proposed framework is generalized approach which employs general compression algorithms and multiple queries are optimized at BS. Repetition of data processing or transmission is checked almost completely in this case.

## **2.7 Summary**

In this chapter we presented a review of various information extraction techniques, highlighting the differences between the DBMS Query processing and WSN Query Processing techniques. A discussion on Energy Efficient approaches based on Mobility, Duty Cycling and Database Optimization has also been carried. Different frameworks proposed for Query Processing have also been presented. In the next chapter, a discussion on experimental environment and the proposed 'CATMOS' framework is done.

### **Proposed Framework for Query processing in WSN: Compression at Input with Multi-query Optimization at Sink (CATMOS)**

A WSN exists or serves its intended purpose as long as its constituent nodes have energy. Hence to extend life of a WSN energy consumption of each node is to be reduced as much as possible. Nodes have to perform sensing, some basic processing and communication through low power radio signals. Communication takes place when a node transmits sensed data in response to a query or in case of occurrence of a predefined event. Data communication gulps a large amount of energy. Therefore, attempts to decrease communication traffic<sup>1</sup> are the most significant steps towards reduction of energy consumption. WSN are being deployed to serve a multitude of applications and each application is has its own metrics of priority.

---

<sup>1</sup>Some work/ contents of this Chapter have been peer reviewed and accepted for publication

[1] Vandana Jindal, Anil Kumar Verma, SeemaBawa, “**Quantitative Reduction in Communication Load for Energy Efficiency in WSN**”, *Wireless Personal Communications (WPC)*, Vol. 83, Number (4), August (II) 2015, pp. 1-16. (SCI-E) (Impact Factor: 0.693)

[2] Vandana Jindal, Anil Kumar Verma, SeemaBawa, “**Green WSN- Optimization of Energy Use through Reduction in Communication Workload**” *International Journal in Foundations of Computer Science & Technology (IJFCST)*, Vol.5, No.3, May 2015, pp. 57-69.

Research work to optimize energy consumption by devising various Medium Access Control (MAC) and routing protocols has taken place and is going on. It has resulted in availability of a number of WSN specific and application specific variants of these protocols. These protocols may serve some applications very efficiently, where as for some applications the same protocols may prove to be unproductive. Since a single large WSN is supposed to serve multiple users for single application or for multiple applications also therefore, a framework has been proposed which employs a generalized approach which can be used for any WSN application. Further gains can be achieved by superimposing application specific protocols by the user. The framework is based on:

- a) Network management by using In network compression;
- b) Query optimization at BS, which merges multiple queries into new queries fewer;
- c) Query language syntax modification by using single syllable Static variables in place of repeated phrases.

All these means have been designed to benefit a WSN irrespective of application being served by it. This general approach will benefit query based applications of data extraction from WSN of any size which isolate user from the underlying network configuration.

### **3.1 Components of the Proposed Framework**

The proposed framework consists of the following components:

#### **3.1.1 WSN Structure**

WSN is made up of a BS/ Sink and numerous sensor nodes. The sensor nodes communicate with the BS in a multi-hop fashion to reduce communication load as multi-hop communication consumes less energy than a single hop for the same distance and route. This is a flat structure for medium size networks.

### 3.1.2 Base station/ Sink

Base station also called a sink is the place for accumulating the sensed data from the sensors in the WSN. It possesses unlimited power supply. Function of the BS is to act as an interface between the network and the user. It injects queries requesting data from the network and receives data sent by nodes as per pre-configured program in nodes or as per instructions from time to time. The queries injected may be continuous, self-generated or manual.

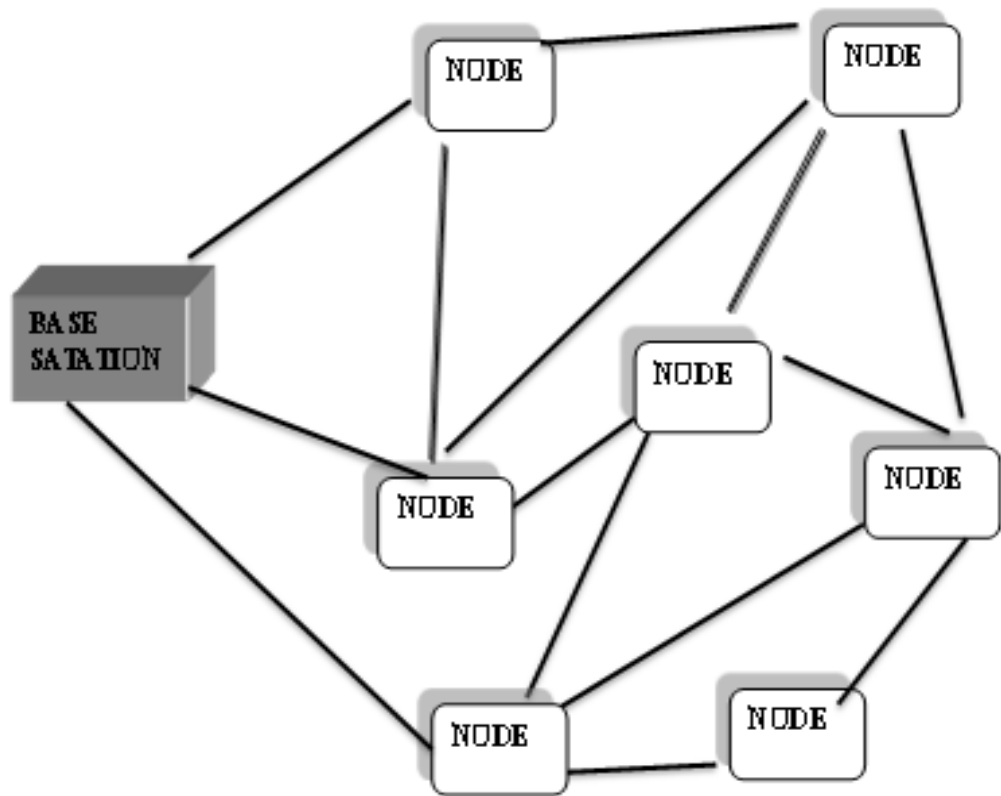


Figure 3.1: A multi hop structure of a WSN

*A continuous query* requires the sensed data periodically.

*A self-generated query* is the one where the node transmits its sensed data to the sink whenever the sensed data fulfills some given condition.

A *manual query* asks for the sensed data from particular sensor(s) randomly as and when desired.

### **3.1.3 Sensor nodes/ Motes**

Sensor nodes or Motes used in the testing are SunSPOT developed by Sun Micro Systems and are built upon the IEEE 802.15.4 standard. The mote is built on Squawk Virtual Machine (VM) [98]. Size of a SPOT is nearly equal to the size of a 3×5 card. It consists of 32-bit ARM9 CPU, 1 MB RAM and 8 MB of Flash memory, a 2.4 GHz radio and a USB interface.

The network platform of SPOT has built-in sensors along with the capability of interfacing with external devices. It is also known as a free range SPOT having a battery, processor board, a sensor board and a sunroof. It houses light and temperature sensors.

### **3.1.4 User**

A user is a person desirous of retrieving information from WSN sitting in a lab/ home/ workplace. User injects a query or number of queries from the PC or any other interface with the WSN. It may be a single user with multiple queries or multiple users with similar query or both.

### **3.1.5 Software**

Programs written in Java[99], utilized for the data reduction are compression algorithms employed at nodes in the network and the proposed query merging/ optimizing algorithm at the sink.

### **3.1.6 SPOT nodes and Software Tools**

Initially sensor nodes used TinyOS as Operating System (OS) and programming was done in NesC. These devices were made by Crossbow Technologies. Sun Microsystems have come up with Java based devices known as SPOTs. These devices have light, temperature and acceleration sensors. The devices need no OS and run Squawk Java

Virtual Machine which serves the purpose of OS and application platform. Each SPOT has a unique ID of sixteen digits. First eight hexadecimal numbers are fixed for each SPOT. Last eight digits fix the identity of each device. Java used is based on J2ME. SunSPOT software can be managed by two tools SPOT Manager and Solarium.

SPOT Manager manages the SunSPOT SDK software. SPOTs can be accessed and configured through this tool. Various Tabs are available to carry out different operations like accessing, configuring, setting up Over TheAir(OTA) (wireless) communications. Any Software Development Kit (SDK) available at Sun\SPOT website can be installed on SPOT.

*Solarium Tool:* It is a host workstation tool to manage application software on SunSPOTs. It can be started from SPOT Manager Tab or from SPOT project directory. Solarium finds all the SPOTs available within radio range. It shows all the identified SPOTs through their numeric address. In Standard Java Micro Edition, only one application can be run at a time in Java VM, that application may have many threads. Squawk enables a SPOT to run many applications on it. It uses a special Isolate Class which prevents one application interfering with other.

Radio connection on a specific port number is given to isolates on first come first basis. All other resources on a SPOT can be shared by Isolates. Master isolate controls the SPOT at system level.

A SPOT application consists of single MIDLET that runs in the master isolate. Solarium is used to start and stop multiple MIDLETs on a SPOT. Several MIDLETs can be packed in a jar file. Virtual SPOTs can also be created in Solarium. Battery of a SunSPOT is 720mAH. Its LED power consumption can be controlled through brightness control. ant.info command is used to get 16 digit unique MAC address of a SPOT. OTA command server can be enabled or disabled.

*Radio Communication:* Radio stream connection interface available in a SPOT provides peer to peer radio protocol with reliable buffered stream based I/O between two devices. Single hop or multiple hop communication can be set up. Radiogram protocol allows exchange of packets between two devices. It supports broadcast mode where there is no

guarantee of delivery. Important data should be uni-cast via radiograms because this mode uses Ack/ retry mechanism whereas broadcast mode does not use this feature.

In case broadcast is to be used packet size must be limited to the extent that each broadcast is maximum three fragments. Single packet broadcast does not require fragmentation header so more space is available for data therefore it increases efficiency.

Solarium can be used to discover SPOT in a WSN and it can manage complete life cycle of applications running on the discovered SPOTs.

### **3.2 Query processor architecture**

Proposed framework is basically based on query type data extraction technique. In WSN, limiting factor for query processing is power consumption so distributed database of WSN is to be managed and protocols for Transport and Network layers are to be optimized. Basic query processing architecture for WSN has:

- a) *Application layer* is at BS, it provides user interface to query the sensor network.
- b) *Database layer* receives queries from the application layer. At BS, database layer contains parser, network catalog manager, query optimizer, buffer manager and scheduler. All these components work for optimizing resource usage by generating query plans, managing memory and systemizing various operations.
- c) *Transport layer* for BS needs adaptation to deal with both TCP for communication with PC and transport protocol used in WSN. Nodes are not addressable so TCP is not used in the WSN.
- d) *Network layer* provides efficient data centric routing algorithms for optimization based on available power with the nodes, energy consumption and hop count for the route chosen. It also takes into account current status of network and detailed query execution plan.

### 3.3 Working of CATMOS

The query injected from the PC is sent to the BS. In response to the query, query optimization is done using query merging technique through the proposed algorithm. BS may encounter multiple queries simultaneously. Response to these queries from the network is optimized by the proposed scheme.

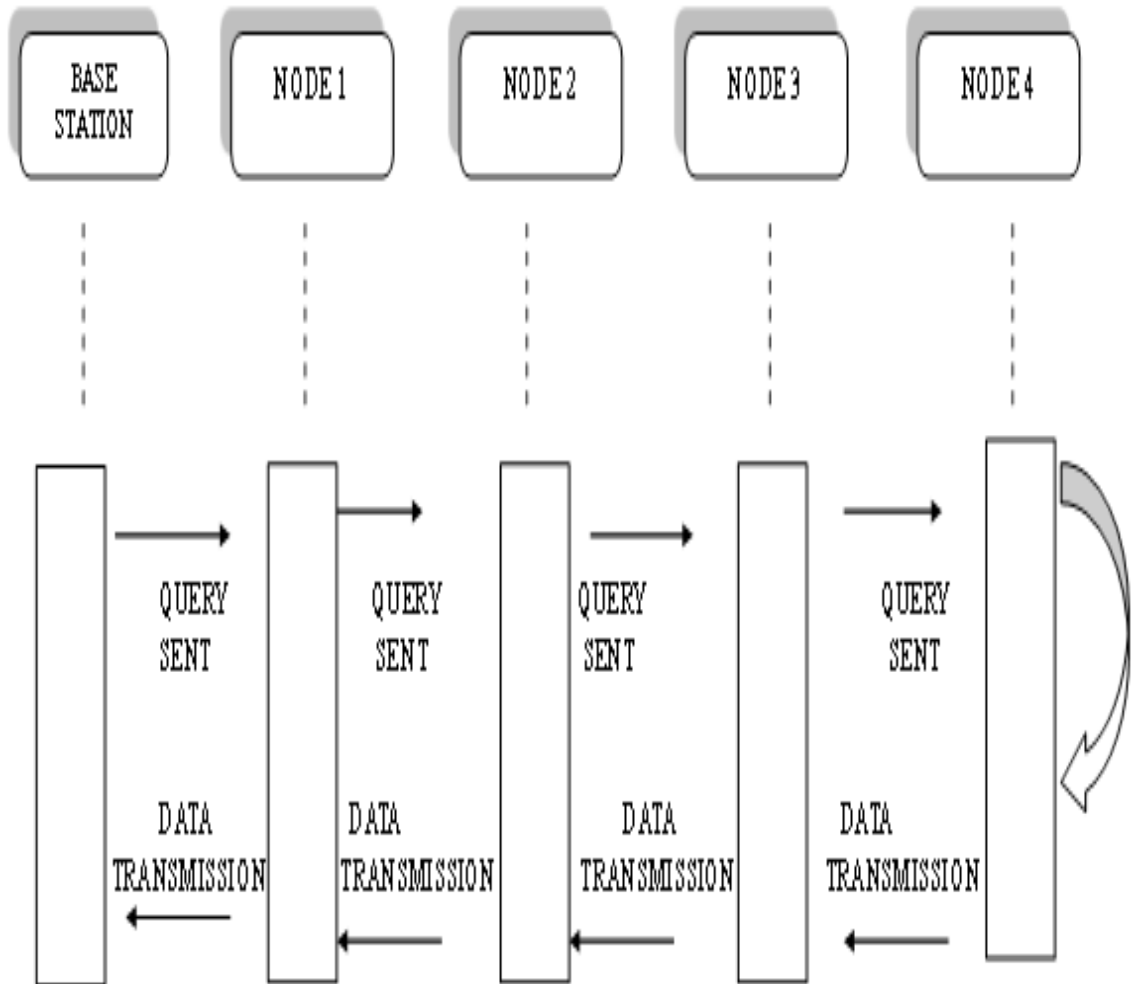


Figure 3.2: Sequence diagram of Query Request

A query is injected into the WSN via the BS whenever some sensed data from the environment is to be sought. The sensor nodes as per the query plan receive the query request from the BS. The sensed data is transmitted to the BS hop-by-hop as per the preconfigured or application specific route.

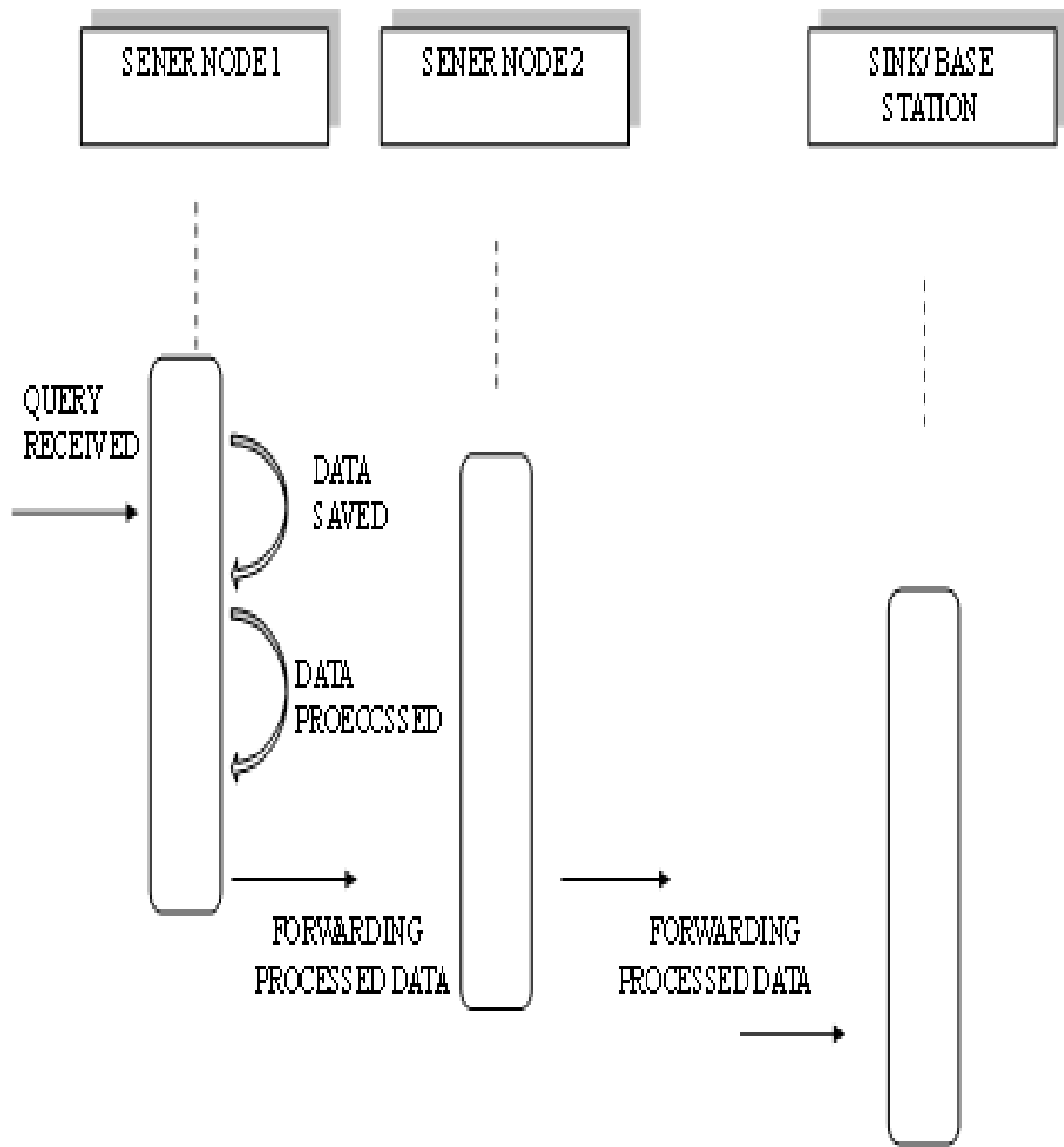


Figure 3.3 Sequence Diagram of sensed data collection

- i. Sensor nodes collect the sensed data, process it i.e. In-network aggregation/ processing and finally transmit the processed data to the next node as per routing protocol which is decided by the requirement of application.
- ii. The sensor node that receives the processed data in turn forwards it to the next node till the processed data reaches the BS.

- iii. The BS on receiving this transmitted data from the sensor nodes makes use of this data for further processing it to convert it into reply of the query or queries running or fresh injected query into the base station.

Sensor data generated by sensing activity of the nodes lies in a distributed manner in the network. This distributed database having WSN specific characteristics is handled by the proposed framework in a judicious and generalized way.

Low complexity small size data compression algorithms compatible to limited memory and processing power of nodes have been employed to eliminate redundancies and compress the data without any loss of useful information.

In-network compression and modification in Query syntax help in reducing the size of data to be transmitted from node to node and nodes to base station and Query optimization at the base-station (sink node) reduces the number of queries injected into the network and thereby reduces the number of transmissions for answering the queries.

In the latest WSN, advanced In-network processing is possible due to increase in the computing capacity of sensor nodes. There are two stages in our optimization process. In Stage-I, the data stream is subjected to compression in the network and in Stage-II Query merging and use of static variable in query syntax is employed at base-station.

BS acts as an interface between the base station application running on the host (PC with Windows platform) and those running on the targets. Energy source available at the base-station is replenishable. In response to a query injected at the base-station query optimization is done using query merging technique. BS encounters multiple queries simultaneously. Response to these queries from the network is optimized. The optimization of multiple queries at the BS helps in achieving significant reduction in energy consumption. As the BS is not resource constrained therefore, there is no limitation on size of optimization algorithm or other tools. Queries posed to WSN generally have similar nature as the main applications are of monitoring be it physical environment, health care, military surveillance. Therefore, these queries may contain identical expressions.

The identical expressions have been replaced with single character called ‘Static Variables’. This reduces the data injected into the network. In addition to this action at sink or BS a query initiates following action at the BS:

When a new query  $q_n$  arrives at a BS where results of a query set  $Q_s$  are already available it is checked whether the query can be satisfied with the available data. If no, the new query is merged with the existing queries and a synthetic query emerges out of this merger. An algorithm has been devised which compares energy benefits of solving the new query as it is or by merging it with already existing queries and generating new synthetic query and finding its solution.

A ‘Gain’ metric has been developed to compare the Gain of query merging. In case, Gain is positive synthetic query is used. The process continues and slowly existing query set transforms into a set of synthetic queries. Once requirement of any query is fulfilled it is removed from the query set. In addition to this as the queries injected into the network have identical expressions to a large extent these identical expressions are replaced with single characters called “*Static Variables*” which further compresses the data to be transmitted in the network.

The proposed framework utilizing “*Compression At input with Multi-query Optimization at Sink*” (CATMOS) runs on Java VM platform. In this work we depict how CATMOS compresses the query syntax and the sensed data along with query merging at the base-station before presenting the final results in response to the user’s queries.

### **3.3.1 Proposed Strategy for Energy Optimization**

Sensor data collected with the help of densely deployed sensors which is to be used by users for useful information extraction is retrieved in an energy efficient manner through the two pronged strategy as detailed below:

### *In-network compression*

The sensed data by the nodes is in the form of continuous data streams. These data streams are tokenized, generating strings of data. These strings are subjected to various compression algorithms such as: Huffman, LZW and Deflate algorithm. These algorithms have been employed to remove redundancies in the sensor data. No useful information has been sacrificed thereby maintaining full data fidelity.

Besides the In-network compression results, which have been evaluated on simulator and are presented in the next section, base-station optimization as described below is also employed.

### *Base station Optimization*

Due to generation and storage of data continuously by nodes in a sensor network a WSN may be considered similar to a distributed database [50, 4]. Assumption that the data is distributed in a database makes the usage of data more comfortable as modification of data becomes easy. Queries are injected to extract information from this database.

Queries may be categorized as one shot queries and continuous queries. Queries reporting the current data only once are termed as one shot queries where as continuous queries are those where the sensors produce and report the data periodically. Multiple queries being studied may be of any of these types.

Optimization of multiple queries at BS by query merging results in reduction of energy consumption significantly. The queries are considered to be long running ones. In case of

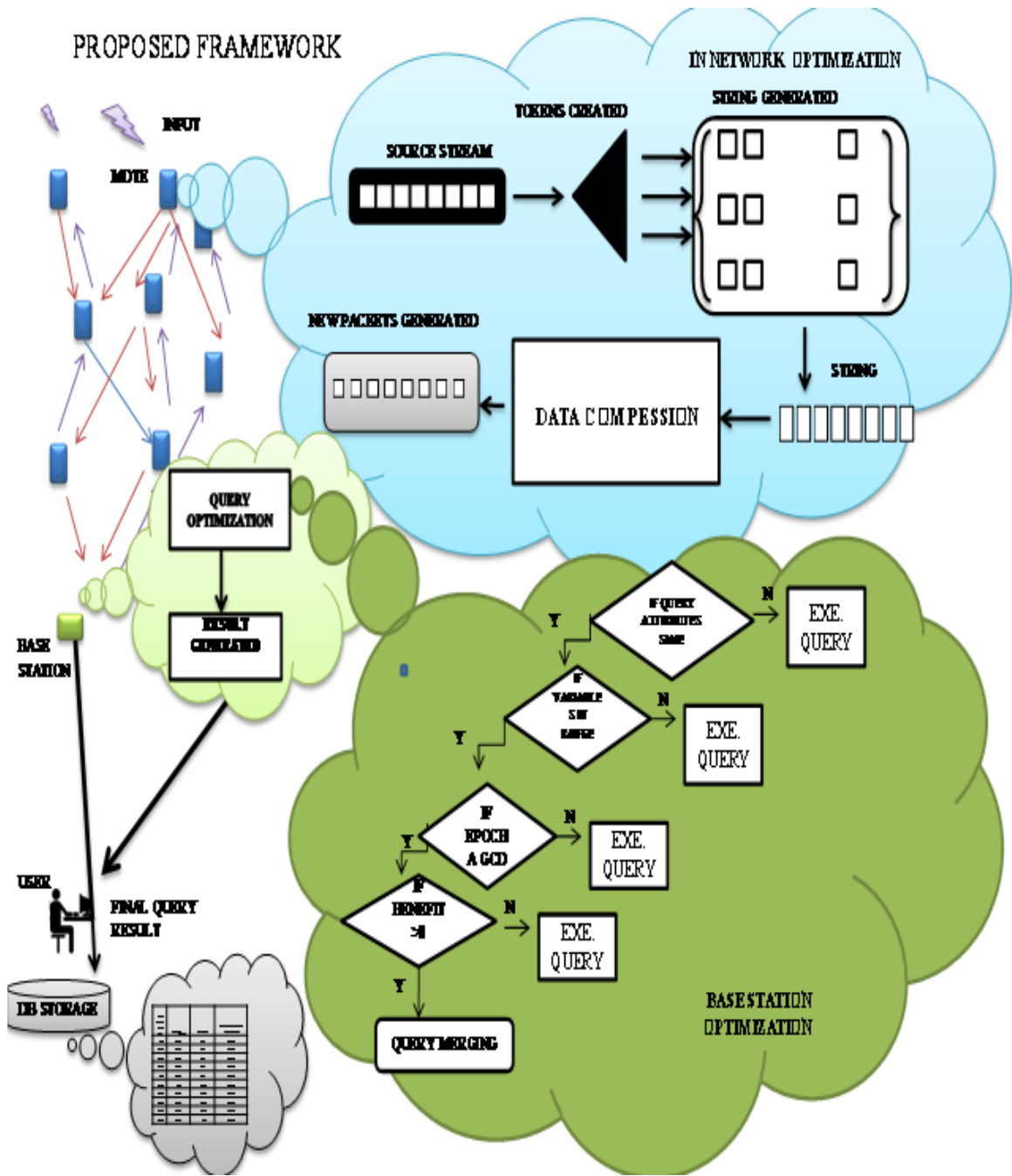


Figure 3.4: The proposed architecture “CATMOS”

long running queries after sometime mostly result dissemination messages are left in the network so a metric in terms of number of result dissemination messages in a unit time is considered to calculate or arrive at the cost involved in answering a query.

Table 3.5: Table of Notations

Symbol	Definition
$Q=\{q_1, q_2, q_3, \dots, q_n\}$	Set consisting of already running queries
$q'_{new}$	a rewritten query of $q_{new}$ defined over $q_1, q_2, \dots, q_n$
<b>ep</b>	Epoch or sample period
<b>sel(p)</b>	Selectivity of a predicate
<b>SC(q)</b>	Selection Criteria of q
$W_{q_i}$	Cost of data acquisition query
<b>Gain</b>	Saved cost in query rewriting
$A_n$	Attributes like temperature, pressure etc.
<b>A(q)</b>	Set of attributes listed
$min_i$	Minimum range of attribute
$max_i$	Maximum range of attribute
$uppervalue_i$	Upper value of predicate p
$lowervalue_i$	lower value of predicate p
$q_1, q_2, q_3, \dots, q_n,$	Queries
$q'$	Synthetic query
$Q_s$	Synthetic query set
<b>GCD</b>	Greatest Common Divisor
$d_{a1}, d_{a2}, \dots, d_{an}$	Decomposition of query
<b>n</b>	Sensor node

*Gain* in cost reduction can be evaluated by finding the difference between the energy cost of newly injected query or queries and query merger. If we assume that each sensor node then has equal chance of being queried, selectivity of a predicate from the data is  $sel(p)$ , in case of multi hop query with sensor at depth  $d$ , cost of data acquisition query ( $W_{q_i}$ ) shall be:

$$W_{q_i} = \frac{sel(p).d}{S_i} \dots\dots\dots(i)$$

where  $S_i$  is sampling rate of the query.  $sel(p)$  depends on the distribution of attributes. Selectivity is computed over the whole data range. If a system has  $A$  attributes for querying and each attribute has a range which falls between  $[min_i, max_i]$  ( $i=1, 2, \dots, A$ ) the selection criteria of the predicate  $p$  having the range  $[upper\ value_i, lower\ value_i]$  may be written as:

$$Sel(p) = \frac{\sum_{i=1}^A (uppervalue_i - lowervalue_i)}{max_i - min_i} \dots\dots\dots(ii)$$

If an existing query set  $Q_1$  is synthesized into new query set  $Q_2$  generated as a result of query merger called synthetic query the cost difference is given as:

$$Cost\ difference = W_{Q_1} - W_{Q_2} \dots\dots\dots(iii)$$

where,  $W_{Q_1}$  is cost of all the queries in  $Q_1$  and  $W_{Q_2}$  is the cost of queries in new set  $Q_2$ . Value must be positive for synthesis to be beneficial. While synthesizing queries into new one it is to be ensured that new query is super set of queries being merged. Semantic correctness is also essential in case of data aggregation queries. Positive difference is denoted by a metric called *Gain*. Gain metric quantifies the saved cost in query rewriting. If we merge two queries  $q_1$  and  $q_2$  into one synthetic query  $q'$ , it should be such that all the data requested by  $q_1$  and  $q_2$  must be requested by  $q'$  also.

$$Gain_{12} = sel(p_1)/s_1 + sel(p_2)/s_2 - sel(p_1 \cup p_2)/GCD(s_1, s_2)$$

We shall write  $q_1$  and  $q_2$  into  $q'$  if and only if  $Gain_{12} > 0$ .

*Gain*<sub>12</sub> > 0 only if

$$\boxed{GCD(s_1, s_2) = s_1 \text{ or } GCD(s_1, s_2) = s_2} \dots\dots\dots \text{(iv) [84]}$$

(‘s’ refers to the sample period or the ‘epoch’)

### 3.3.1.1 Mathematical Analysis of the proposed approach

Dual strategy of In-network processing/ aggregation and Base station optimization are discussed in details in this section.

*In-network aggregation:* This involves data compression at the node level. In response to a query injected into the network by the base station, the data is sensed by the sensors and is stored in the nodes in the shape of data streams. The data stream is converted into a string format and is subjected to various compression algorithms. The data is then transmitted to the Base Station (BS) where it’s optimization takes place.

The main objective is to reduce the data size by eliminating spatio temporal redundancies. Data fidelity is given importance and there is quantitative reduction in data only without any reduction in quality. Reduction in data to be communicated, results into less draining of scarce energy of nodes. Gains through compression have been verified on a simulator. Brief simulator setup & results have been discussed in the following sections.

Whenever a query is injected into the network it drives free nodes of the network into action and they sense continuous stream of data from the environment. The queries that flow into the WSN after initiation from the base station generally have many identical expressions. In order to reduce communication load these identical expressions are replaced by a single term. It acts as a sort of compression at the time of compilation itself.

The base station is connected to a PC where power is not the limitation. The results generated by the already running queries within the network are explored to obtain the reply for newly injected queries before floating them into the network. A modern WSN is able to perform multiple tasks for multiple users as it is versatile enough to satisfy

information needs of widely varying applications simultaneously. Hence it is flooded with multiple queries at any instance. Here Query merging which is the process to rewrite the queries with the help of already running queries is resorted to. For this, query rewriting algorithm has been proposed which merges new queries with various already running queries. *Rewriting the Query* is demonstrated as follows:

$q_{inject1} = \text{select nodeid, illumination from sensors where } 40 < \text{illumination} < 80 \text{ epoch } 3\text{sec}$

$q_{inject2} = \text{select nodeid, illumination from sensors where } 120 < \text{illumination} < 160 \text{ epoch } 6\text{sec}$

Now a new query is written such that the data requested by new query should be obtainable from the results of the two queries. In this case attributes and predicates of the new query will be Union of these two queries and epoch duration will be the Greatest Common Divisor (GCD) of the epoch of these two queries. Therefore new query  $q'$  should be:

$q' = \text{select nodeid, illumination from sensors where } 40 < \text{illumination} < 160 \text{ epoch } 3\text{sec}$

In this way by using Greedy Algorithm one can merge two or more queries. The algorithm calculates the Gain from the merger. If Gain is more than zero then we go ahead with the merger.

Now let us assume these queries  $q_1, q_2, q_3$  such as:

$q_1 = \text{"select nodeid from sensors where } 150 < \text{pressure} < 450 \text{ epoch } 2\text{sec"}$

$q_2 = \text{"select nodeid from sensors where } 40 < \text{pressure} < 360 \text{ epoch } 4\text{sec"}$

$q_3 = \text{"select nodeid from sensors where } 20 < \text{pressure} < 320 \text{ epoch } 4\text{sec"}$

Here,  $\text{Gain}(q_1, q_2) = 300/2 + 320/4 - 410/2$

$$= 150 + 80 - 205$$

$$= 230 - 205 > 0$$

$\text{Gain} > 0$ , therefore queries will be merged. Hence the new queries may be written as:

$q'_1 = \text{select nodeid from sensors where } 40 < \text{pressure} < 450 \text{ epoch } 2\text{sec}$

$\text{Gain}(q'_1, q_3) = 410/2 + 300/4 - 430/2$

$$= 210 + 75 - 215 > 0$$

Query will be merged  $q'_2$  will be executed.

$q'_2 = \text{select nodeid from sensors where } 20 < \text{pressure} < 450 \text{ epoch } 2\text{sec}$

Similarly, we can explore the possibility whether we can answer a new query by reusing the results of existing queries. We can save unnecessary data communication and save energy. For example:

If the existing running queries in a network are:

$q_1 = \text{select nodeid from sensors where illumination} > 25 \text{ epoch } 2\text{sec}$

$q_2 = \text{select nodeid from sensors where illumination} > 60 \text{ epoch } 4\text{sec}$

If a new query is injected to the BS as follows:

$q_{\text{new}} = \text{select nodeid from sensors where illumination} > 40 \text{ and humidity} > 70 \text{ epoch } 8 \text{ sec}$

In this case we can answer the new query from the results of the existing queries  $q_1$  and  $q_2$  without injecting the query into the network.

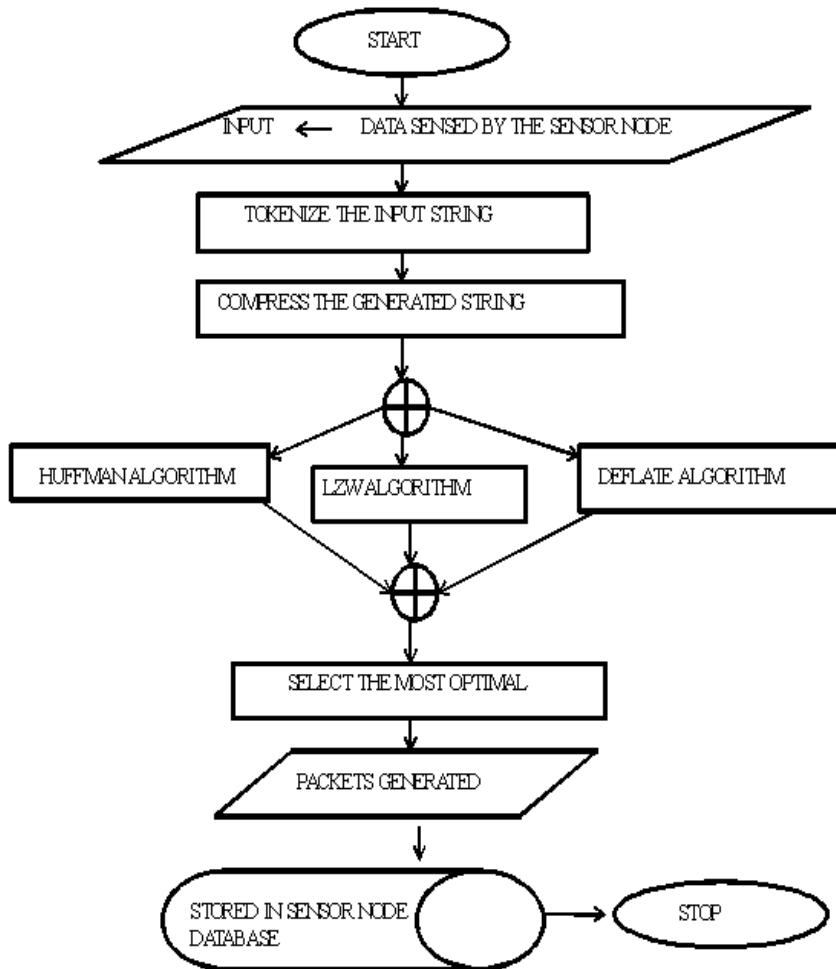


Figure 3.6: In-network processing

In case  $\text{Gain} < 0$ , new queries are not integrated and go directly into the existing synthetic query list. If  $\text{Gain} > 0$ , queries are synthesized, new synthetic query is then checked for positive Gain with other queries of the set. The iterative algorithm is so designed that any achievable positive Gain is exploited fully.

---

*Pseudo code/ Algorithm for In-network processing*

---

Step 1: Start

Step 2: Input: Data sensed by the sensor node.

Step 3:  $d \rightarrow \{d_1, d_2, d_3, d_4, d_5, \dots, d_n\}$  /\*Data = d, sensor node = n, queue = Q\*/.

Step 4: Compressing the data by using three different algorithms.

Step 4.1 Huffman Algorithm

```

huffman(d) {
  n1 = |d|;
  Q = d;
  for (i= 1 to n-1){ z= new node;
    left[z] = extract - min(Q);
    right[z] = extract - min(Q);
    f[z] = f[ left[z]] + f[right[z]];
    insert(Q, z);}

  return extract-min(Q)}

```

Step 4.2 LZW Algorithm

```

w = nil;
while( read a character d)
{ if wd exists in the dictionary
  w=wk;
else
  add wk to the dictionary;
}

```

```
output the code for w;  
w=d;  
}
```

#### Step 4.3 Deflate algorithm

Step4.3.1: set the coding position to the beginning of input stream

Step4.3.2: if coding position is not at the end of input stream, search the window for the longest match with the look ahead buffer;

else algorithm terminates.

Step4.3.3: if find match,

output (off, length, d),dc is the character following the match,  
coding position and window move length+1 bytes forward;  
else goto step4.3.4.

step4.3.3: output current character at coding position, coding position and windows move 1 byte forward;

goto step 4.3.2.

Step 5: select the most optimal compression from these three algorithms.

Step 6: Generate the packets from the compressed data.

Step 7: Generate packets stored in sensor node database.

Step 8: Stop.

*Synthesis of Multiple Queries:* The section here presents multi-query optimization algorithms. The base-station is the interface between the network and the user. User sends queries and obtains the result at the base station. Base station is not resource constrained as the nodes. Therefore, base-station is used to filter out the redundant load of multiple queries into the network. Multi-query optimization algorithm rewrites a set of similar queries into a new set of queries before injecting them into the network, so that redundant data requests can be eliminated as much as possible. Correctness of semantics

of queries is to be ensured while rewriting new query. All this is achieved with the help of different algorithm designed to obtain maximum Gain.

*Treatment of a new query:* If a new query  $q_n$  arrives at a base station where results of a synthetic query set  $Q_s$  are already being calculated or obtained from the network, the algorithm will evaluate the benefit of rewriting the new query with the existing synthetic queries and find the most beneficial (in terms of cost) one. A new query is generated by merging the new query and existing query where the Gain is maximum. If there is no such query then new query is directly added to the set  $q_s$ . Iterative evaluation of cost reduction by integration of new synthetic query in the  $q_s$ : When a synthetic query is generated on arrival of new query as explained above the previously existing query set  $Q_s$  gets modified.

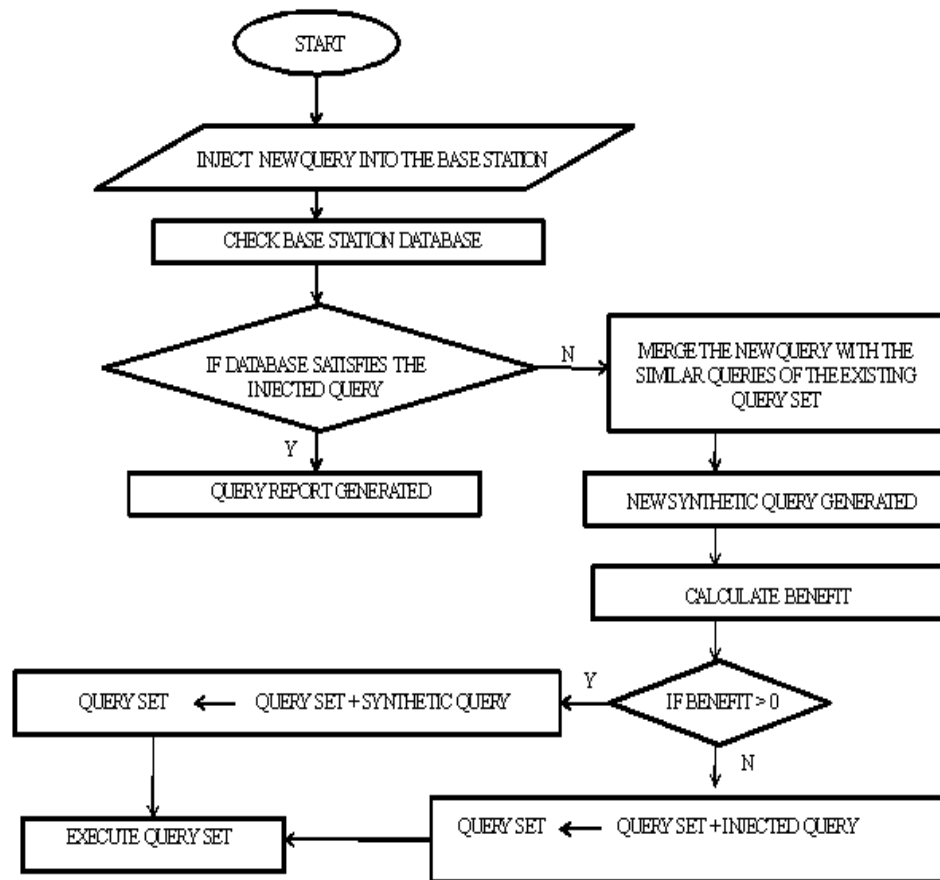


Figure 3.7: Base station optimization

It is evaluated through successive iterations through algorithm whether new synthetic query can be synthesized with any of the existing queries in  $Q_s$  to achieve further Gain. If so the pair is rewritten and the new one is again checked. This iterative process is continued till no further Gain is achievable. When duration of a new injected query gets finished it is to be removed from the query set being evaluated. Query list is updated again to the previously optimized status prevailing before injection of the query. In this way Multiple Query Optimization (MQO) design should be scalable-every new query is treated individually for positive Gain, energy efficient and adaptable-no redundant query should be there in the system.

*Query Rewriting* is employed for answering the newly injected queries by reusing the results of the already existing queries. This results into optimization of resource usage as duplicate data requests can be removed. The notations used in the algorithm are:

$Q = \{q_1, q_2, \dots, q_n\}$  the set consisting of already running queries;

$a_n$  = The attributes like temperature, pressure etc.;

$q'_{new}$  = a rewritten query of  $q_{new}$  defined over  $q_1, q_2, \dots, q_n$ ;

EP = Epoch is the time interval of taking the readings; sampling rate;

SC(q) = Selection Criteria of q;

A(q) = Set of attributes listed;

S(q) = Set of attributes that are in selection criteria SC;

(sc=  $p_1 \wedge p_2 \dots$ ) e.g.sc= (light > 100)  $\wedge$  (temp<30)

Decomposition of query  $q_{new}$  into  $d_{a_1}, d_{a_2}, \dots, d_{a_n}$ ;

$$d_{a_i} = \begin{cases} \prod_{nodeid} \sigma_{sc}(q_{new}, a_i)(sensors), & \text{if } a_i \in A(q_{new}) \\ \prod_{nodeid} \sigma_{sc}(q_{new}, a_i)(sensors), & \text{if } a_i \in S(q_{new}) - A(q_{new}) \end{cases}$$

for  $1 \leq i \leq n$ .

e.g.  $d_{temp} = \prod_{nodeid, temp} \sigma_{temp}(q_{temp<30})(sensors)$

$d_{light} \prod_{nodeid} \sigma_{light<100}(sensors)$

---

*Algorithm for Rewriting the Query*

---

```

1. Let  $A(q_{new}) \cup S(q_{new}) = \{a_i\}$ ;
   where  $i = 1, 2, 3, \dots, n$ ;          /* composing the Prospect Query set*/
2.  $Q' = Q$ ;
3. for  $j=1$  to  $m$ 
   |   for  $q_j \in Q'_j$  do
   |   |   if  $((EP(q_{new}) \% EP(q)) \neq 0) \wedge (SC(q) \wedge SC(q_{new}) == \text{false})$ 
   |   |   |   remove  $q$  from  $Q'$ ;
   |   |   |   end
   |   |   end
   |   end
end
4. for  $i=0$  to  $n$ ,  $A(q_{new}) \cup S(q_{new})$ 
   |   do  $Q' a_i = \{q \mid (q \in Q') \wedge (a_i \in P(q))\}$ ;
   |   end
end
5. for each  $Q' a_i$  ( $1 \leq i \leq n$ )
   |   do compute test  $C(q_{new}, a_i) \rightarrow \forall_{q \in Q' a_i} SC(q)$ 
   |   |   if test == true
   |   |   |   hold;
   |   |   |   else return;
   |   |   end
end
6. for  $i=1$  to  $n$  in  $A(q_{new})$ 
   |   do  $d_{ai} = \bigcup_{q \in Q' a_i} \prod_{\text{nodeid}} a_i \sigma_{sc}(q_{new}, a_i)(q)$ ;
   |   end
end
7. for  $i=1$  to  $n$ 
   |    $S(q_{new}) - A(q_{new})$ 
   |   |    $d_{ai} = \bigcup_{q \in Q' a_i} \prod_{\text{nodeid}} (\sigma_{sc}(q_{new}, a_i)(q))$ ;
   |   end
end
8. Return  $q'_{new} = d_{a1} a_2 \dots d_{an}$ ;

```

In this chapter we have proposed an energy efficient In-network compression technique to reduce traffic in WSN facilitating an extended life time of the network. At the base-station query re-writing method is used for writing a new query if possible from the already injected queries. Both methods result into reduced energy consumption of the nodes. The work is implemented using the SPOT wireless platform. After the data has been sensed by the sensor of a free-range SPOT that is installed in the network as a sensor node, the data compression algorithm comes into play. This compressed data is then transmitted to the BS SPOT which in turn sends the data to a PC. Results indicate that compression rate and power reduction are quite impressive. The query optimization processing algorithm described in this chapter results in cutting the cost in terms of energy and bandwidth usage, both of which are scarce resources in a WSN.

### **3.5 Summary**

The chapter presents in details about the proposed Framework ‘CATMOS’, explaining all its components, the software employed for its working, and query processing architecture. Query optimization processing algorithm was discussed for rewriting of query at the base station.

## Chapter-4

### Simulation Setup, Results and Analysis

The goal of this simulation study is the evaluation of the achievable energy gains when data compression is applied prior to packet transmission. We examine a multi-hop scenario, where a node periodically delivers the data to the sink (BS). The data collected through high density distributed WSN are immense. In applications like temperature monitoring, the data collected at the free nodes needs to be transferred to the base-station periodically with low latency so that the data available is up-to-date. When readings are taken at regular time intervals, they are not expected to change significantly, but need to be transferred to the base-station at regular intervals.

#### 4.1 Simulation Setup

SunSPOT mote used in the study<sup>2</sup> is a WSN mote developed by Sun Microsystems. The device is built upon the IEEE 802.15.4 standard. Unlike other mote systems, The SPOT is built on Squawk VM [100]. A SPOT is about the size of a 3×5 card with 32-bit ARM9

---

<sup>2</sup>Some work/ contents of this Chapter have been peer reviewed and accepted for publication

- Vandana Jindal, Anil Kumar Verma, SeemaBawa, “**Two pronged Strategy for Energy Optimization in WSNs by using In-network Compression and Synthesis of Multiple Queries at Base-Station**”, *In Proceedings of the International Conference on Systems, Control, Signal Processing and Informatics (SCSI 2015), Barcelona, Spain, April 7<sup>th</sup> -9<sup>th</sup>, 2015, pp 135-141.*

CPU, 1 MB RAM and 8 MB of Flash memory, a 2.4 GHz radio and a USB interface. The network platform of SPOT has built-in sensors along with the capability of interfacing with external devices. Two kinds of SPOTs i.e. free-range SPOT and base-station SPOT are present. The anatomy of free range SPOT has a battery processor board, a sensor board and a sunroof.

The sensors present are capable of measuring acceleration, temperature and light intensity. The base-station SPOT does not have a sensor board. It acts as an interface between the base station application running on the host (PC with Windows platform) and those running on the targets. The host application is a J2SE program [101] and target application is a Squawk Java Program. Two of the platforms used here are the SPOT Manger tool and Net Beans 7.0 Integrated Development Environment. The SPOT Manger tools are – the SPOT Manager and Solarium [102]. Each SunSPOT has an IEEE network number. Solarium has an emulator for running applications on a virtual SPOT. Initially sensor nodes used TinyOS as operating system and programming was done in NesC.



Figure 4.1: SunSPOTs with an AA battery[107]

The compressed data is then transmitted to the user's end via the base station thus reducing the quantum of data transmission. The reduction in the size of the data leads to saving energy of the nodes in the WSN.

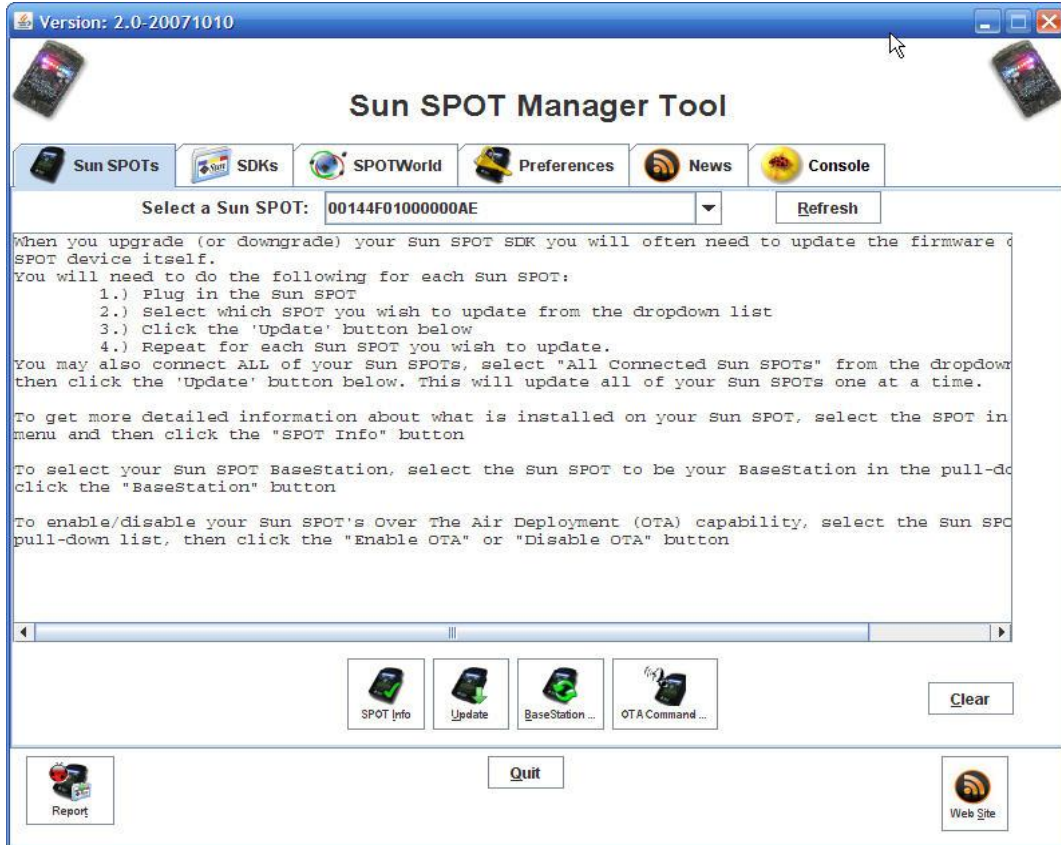


Figure 4.2: The SPOT Manager[107]

NetBeans is a software development platform written in Java. The NetBeans Platform allows applications to be developed from a set of modular software components called modules. Applications based on the NetBeans Platform, including the NetBeans integrated development environment (IDE), can be extended by third party developers[103]. The NetBeans IDE is primarily intended for development in Java, but also supports other languages, in particular PHP, C/C++ and HTML5 [96][10]. NetBeans is cross-platform and runs on Microsoft Windows, Mac OS X, Linux, Solaris and other platforms supporting a compatible JVM.

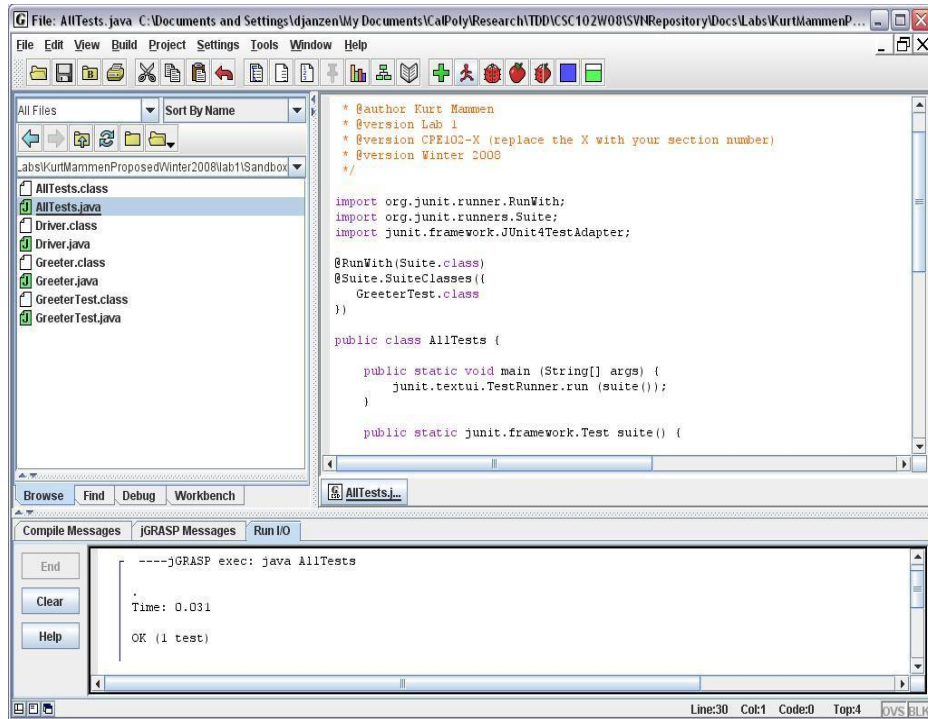


Figure 4.3: NetBeans[107]

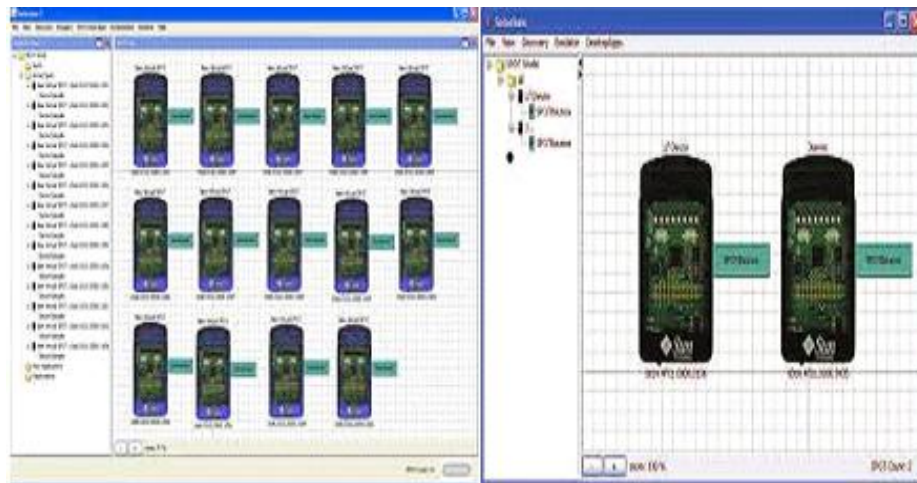


Figure 4.4: Solarium Application[107]

The findings have been summarized in tabular form and then the graphical representations have been depicted.

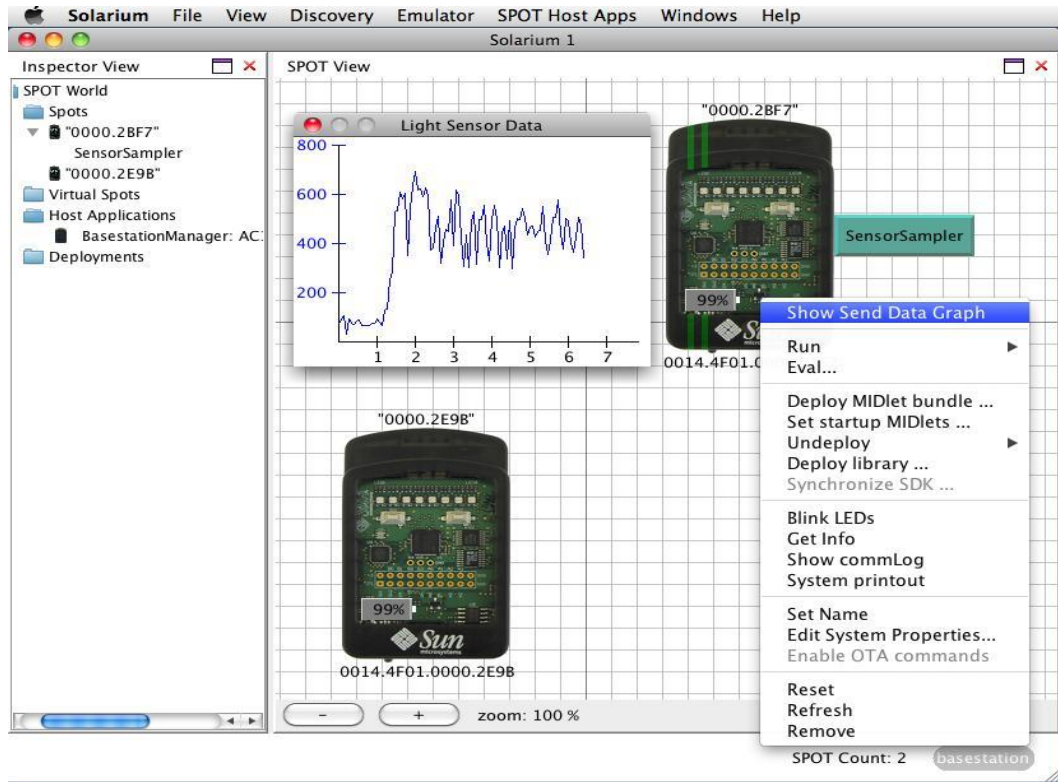


Figure 4.5: Solarium showing the graph from the sensed data[107]

Table 4.6: Raw data (bytes) and Compressed Data (bytes) using Huffman/ LZW/ Deflate algorithms

Range		Epoch	Raw Data (bytes)	Compressed Data using Huffman Algo. (bytes)	Compressed Data using LZW Algo. (bytes)	Compressed Data using Deflate Algo. (bytes)
Temp <sub>(min)</sub>	Temp <sub>(max)</sub>					
10	50	5	296	152	282	360
10	40	8	296	144	282	360
10	50	5	304	152	290	368
10	35	40	304	152	290	368
15	25	50	304	152	290	368
15	30	10	304	152	290	352
5	55	120	312	160	305	376
25	50	130	312	160	291	376
30	45	60	304	152	297	368
35	50	70	312	160	298	376

Table 4.7: Compression Factor (CF) (%age compression using Huffman/ LZW/ Deflate algorithm)

Huffman (CF)	LZW (CF)	Deflate (CF)
51.351351	95.27027	121.62162
48.648649	95.27027	121.62162
50	95.394737	121.05263
50	95.394737	121.05263
50	95.394737	121.05263
50	95.394737	121.05263
50	95.394737	115.78947
51.282051	97.75641	120.51282
50	97.697368	121.05263
51.282051	95.512821	120.51282

Table 4.8: Percentage compressions (Huffman / LZW Compression algorithm)

Huffman (CF)	Huffman (CF with SVar.)	Huffman (Comp. Diff)	LZW (CF)	LZW (CF with SVar.)	LZW (Comp. Diff)
51.351351	42.105263	9.246088	95.270270	95.394736	-0.124466
48.648648	36.842105	11.806543	95.270270	90.789473	4.480796
50.000000	40.000000	10.000000	95.394736	95.625000	-0.230263
50.000000	40.000000	10.000000	95.394736	95.625000	-0.230263
50.000000	40.000000	10.000000	95.394736	95.625000	-0.230263
50.000000	40.000000	10.000000	95.394736	91.250000	4.144736
51.282051	42.857142	8.424908	97.756410	100.000000	-2.243589
51.282051	42.857142	8.424908	93.269230	91.666666	1.602564
50.000000	40.000000	10.000000	97.697368	100.000000	-2.302631
51.282051	38.095238	13.186813	95.512820	95.833333	-0.320512

## 4.2 Test cases, test results and test analysis

### A. Test Case I

#### Test Objectives

Tests were conducted with the objective to find out change in energy consumption before and after In-network processing of raw data streams at nodes using compression algorithms with strict monitoring so that there is no data loss. For this, in Test Case-I the data streams were compressed using the Huffman compression. The test bed used for the test was 'Solarium'. Verification of the output was done employing synthetic data.

#### Test Results

A graph was plotted using raw streams of data and the final results obtained after processing this sensed data. The x-axis denoted the number of queries whereas the y-axis depicted data generated. The plotted graph showed both the results with uncompressed raw data streams and after the streams were compressed.

The final graph plotted between the number of queries executed against their respective data sizes i) without compression and ii) with Huffman Compression are as under:

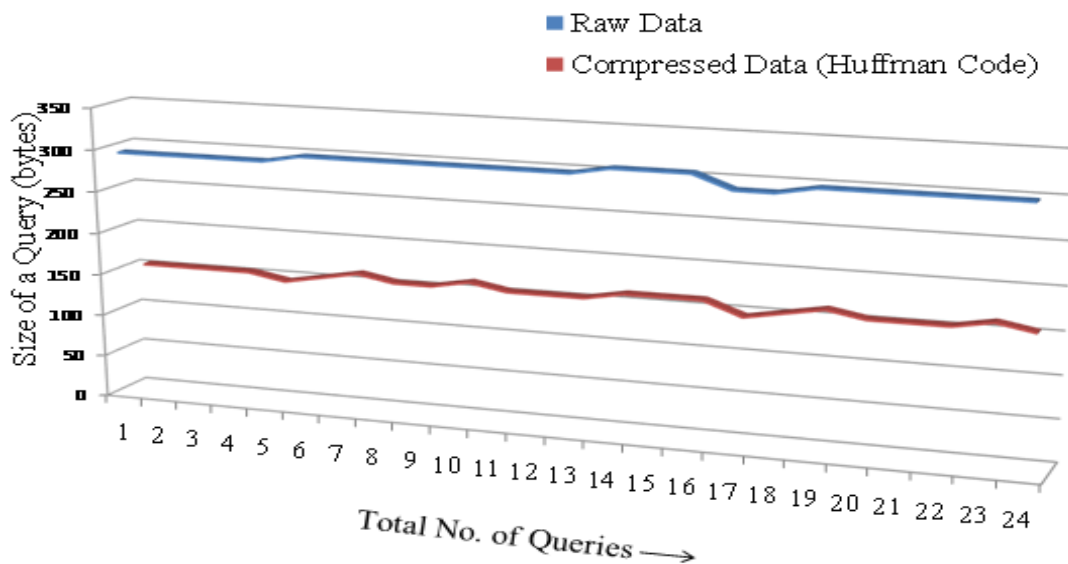


Figure 4.9: Raw Data vs. Compressed Data using Huffman Compression algorithm

### *Test Result Analysis*

The main objective of the study is to assess the achievement in energy savings by using data to compression before transmitting data to the base-station via other nodes as per the established route. Compressing the data leads to decrease in the size of data that is to be transmitted. Reduced data size leads to reduced number of the radio transmissions. Reduction in radio transmissions leads to lower energy consumption. It is observed that the produced data gets reduced through the applied compression algorithm without losing its contents. The compression factor was attainable up to 48.64%.

### **B. Test Case II**

#### *Test Objectives*

Tests were conducted with the objective to find out change in energy consumption before and after In-network processing of raw data streams at nodes using compression algorithms with strict monitoring so that there is no data loss. For this, in Test Case-II the data streams were compressed using the LZW compression. The test bed used for the test was 'Solarium'. Verification of the output was done employing synthetic data.

#### *Test Results*

A graph was plotted using raw streams of data and the final results obtained after processing this sensed data. The x-axis denoted the number of queries where as the y-axis depicted data generated. The plotted graphs showed both the results with uncompressed raw data streams and after the streams were compressed.

The final graph plotted between the number of queries executed against their respective data sizes i) without compression and ii) with LZW compression are as under:

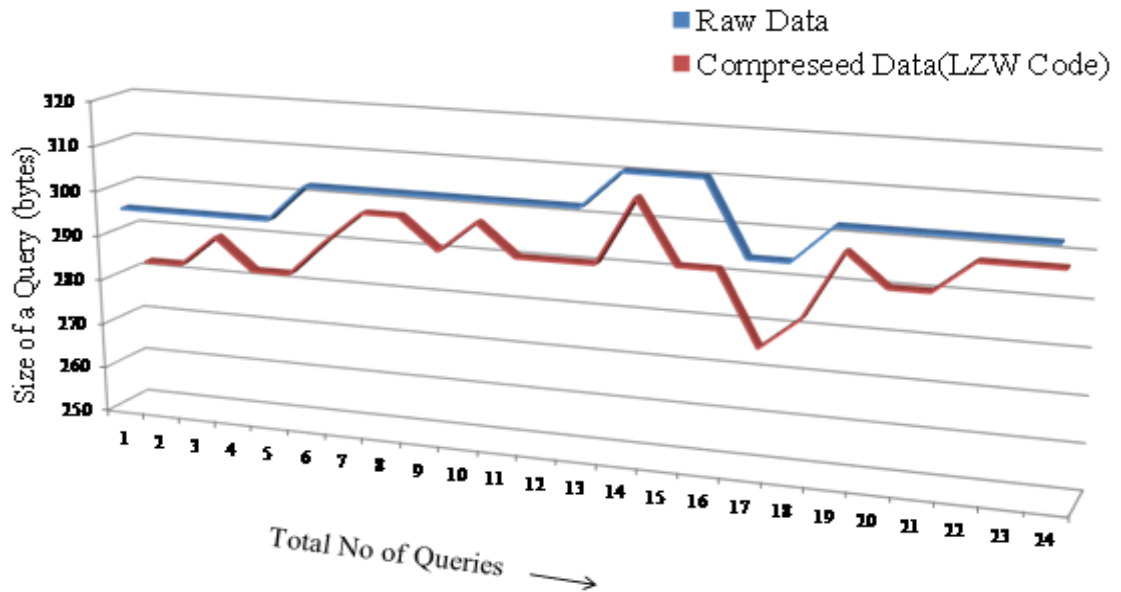


Figure 4.10: Raw Data vs. Compressed Data using LZW algorithm

### *Test Result Analysis*

The main objective of the study is to assess the achievement in energy savings by using data to compression before transmitting data to the base-station via other nodes as per the established route. Compressing the data leads to decrease in the size of data that is to be transmitted. Reduced data size leads to reduced number of the radio transmissions. Reduction in radio transmissions leads to lower energy consumption. It is observed that the produced data gets reduced through the applied compression algorithm without losing its contents. The compression factor was attainable up to 92.90%.

### *C. Test Case III*

#### *Test Objectives*

Tests were conducted with the objective to find out change in energy consumption before and after In-network processing of raw data streams at nodes using compression algorithms with strict monitoring so that there is no data loss. For this, in Test Case-III the

data streams were compressed using the Deflate compression. The test bed used for the test was ‘Solarium’. Verification of the output was done employing synthetic data.

*Test Results*

A graph was plotted using raw streams of data and the final results obtained after processing this sensed data. The x-axis denoted the number of queries where as the y-axis depicted data generated. The plotted graphs showed both the results with uncompressed raw data streams and after the streams were compressed.

The final graph plotted between the number of queries executed against their respective data sizes i) without compression and ii) with Deflate algorithm are as under:

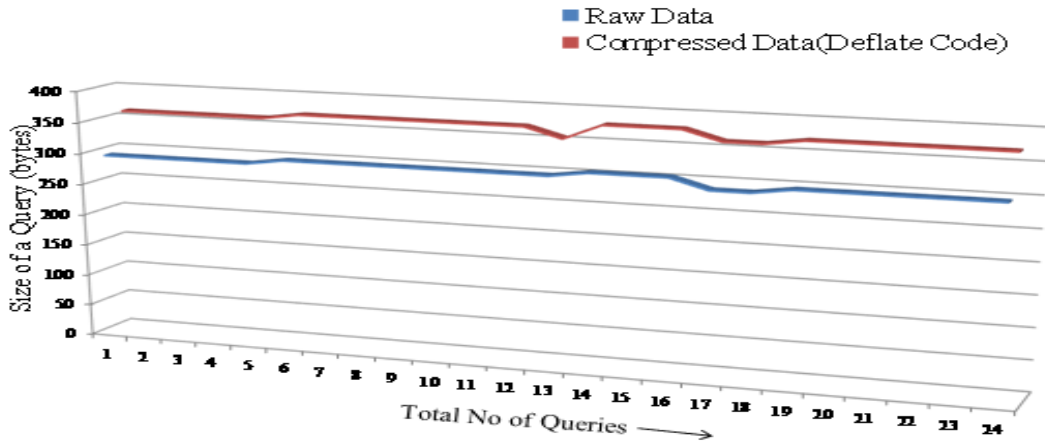


Figure 4.11: Raw Data vs. Compressed Data using Deflate algorithm

*Test Result Analysis*

The main objective of the study is to assess the achievement in energy savings by using data to compression before transmitting data to the base-station via other nodes as per the established route. Compressing the data leads to decrease in the size of data that is to be transmitted. Reduced data size leads to reduced number of the radio transmissions. Reduction in radio transmissions leads to lower energy consumption. It is observed that the produced data gets reduced through the applied compression algorithm without losing its contents. The compression factor was not attainable up to the mark. It was 115.78% (more than the original size) which failed to serve the purpose.

#### D. Test Case IV

##### Test Objectives

The objective of the test case IV is to verify the effect of query merging at base-station data to be communicated in response to a number of queries. The test is executed on the test bed “Solarium”. It is to be established whether the results obtained using the synthetic data fulfils the intent of answering multiple queries with reduced communication load in the network.

##### Test Results

The graph in Figure 4.10 is obtained by employing three techniques i.e. i) injecting autonomous queries, ii) query blending/ merging and iii) rewriting of queries conjointly with the merger of the previously injected queries at the base-station. The values along the x-axis depicts the total number of sensor readings generated and values along the y-axis depict the number of sensor readings required to be relayed through the established route to the sink (base-station).

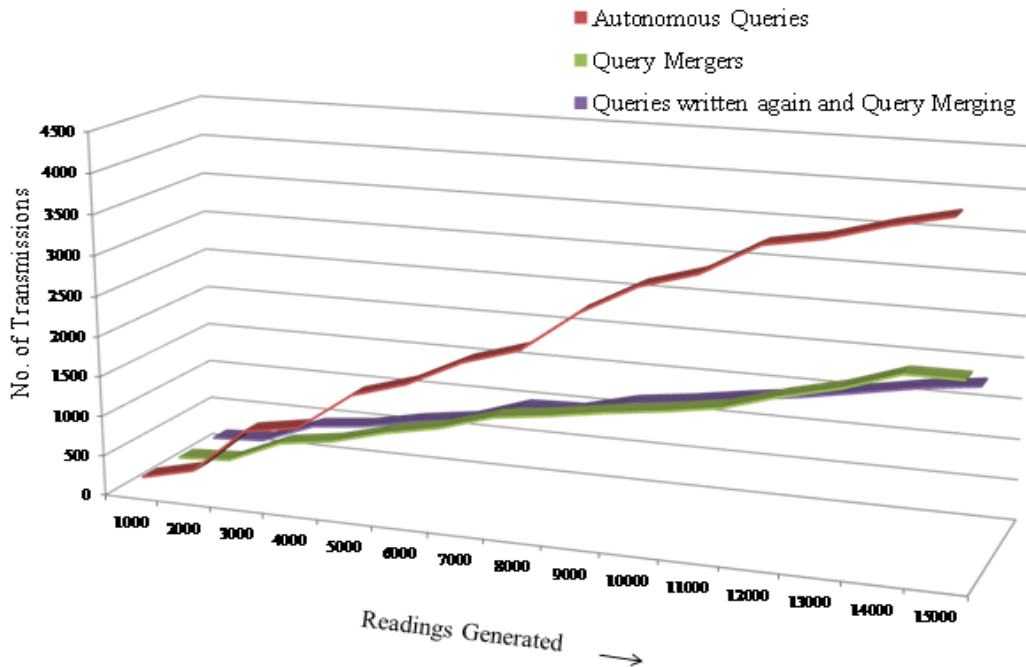


Figure 4.12: Results obtained from Number of Sensor readings transmitted and received

### *Test Result Analysis*

Undoubtedly, the integration or the merger of the queries is beneficial in terms of data reduction. It is established that rewriting queries from amongst the lot in operation as most of the queries are continuous queries with new injected queries proves to be profitable. The technique shows positive results. The total number of bits required to be communicated gets reduced. Out of the two approaches, the latter proved to be more effective in energy conservation as it greatly reduces.

In order to further extend the battery life of the sensors, we made use of ‘Static variables’ i.e. substituting the repeated text in queries by a single variable, which can be thought of compression at the time of compilation. By employing these static variables, the compression level enhanced further.

#### **E. Test Case V**

### *Test Objectives*

The objective of the test is to evaluate the amount of energy savings after employing static variables during compilation. The approach was extended on the queries to be injected from the base station. Huffman compression was employed further on the query to be injected into the WSN. The test bed used for the experiment was ‘Solarium’.

### *Test Results*

A graph was plotted with the x-axis depicting the number of queries where as the y-axis depicting the compressed data with Huffman code without using Static Variable in queries and compressed data with Huffman code and using Static Variable at compilation time.

### *Test Result Analysis*

The main objective of the study was to assess the achievement in energy savings by subjecting queries to static variables during compilation. It is evident that the objective of enhancing energy savings was achieved. There was a further increase in saving up to a maximum of 12.63% after the compression when data subjected to Huffman code.

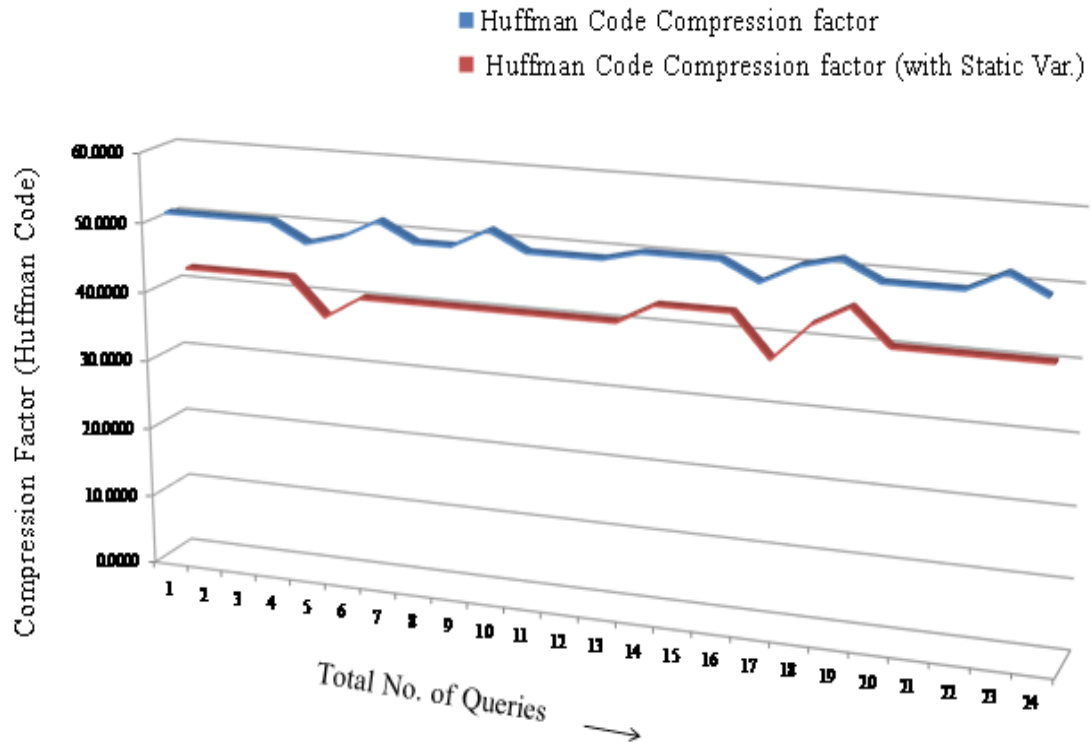


Figure 4.13: Compression Factor (without Static Variable) vs. Compression Factor (with Static Variable) using Huffman Compression Algorithm

**F. Test Case VI**

*Test Objectives*

The objective of the test is to evaluate the amount of energy savings after employing static variables during compilation. The approach was extended on the queries to be injected from the base station. LZW compression was employed further on the query to be injected into the WSN. The test bed used for the experiment was ‘Solarium’.

*Test Results*

A graph was plotted with the x-axis depicting the number of queries where as the y-axis depicting the compressed data with LZW code without using Static Variable in queries and compressed data with LZW code and using Static Variable at compilation time.

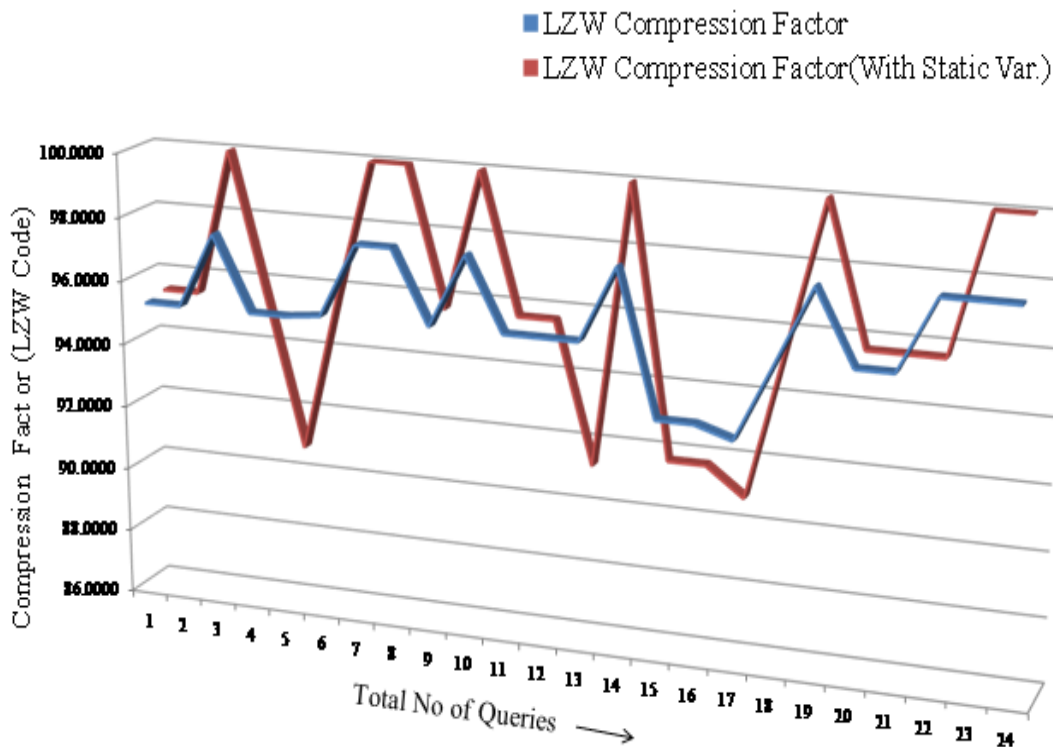


Figure 4.14: Compression Factor (without Static Variable) vs. Compression Factor (with Static Variable) using LZW Algorithm

#### *Test Result Analysis*

The main objective of the study was to assess the achievement in energy savings by subjecting queries to static variables during compilation. It is evident that the objective of enhancing energy savings was achieved. There was a further increase up to maximum of 4.48% after the compression when data subjected to LZW code.

We have proposed use of low power consuming compilation techniques at node level and query merging at base station for optimization of data transmission in order to decrease the energy consumption for serving multiple applications by a WSN.

The application of this optimization resulted in a modest 10.29% reduction in the power consumption. Energy savings in one network may seem to be miniscule but as the number of deployed WSNs is increasing day by day, these savings translate into huge

savings globally. Therefore any saving in energy through software means, results in benefits for the times to come. Reduction in energy consumption and possibility of deployment of WSN in any unimaginable location will go a long way in achieving environmental protection targets goals also.

Table 4.15: Comparative features of various frameworks with CATMOS

	<b>TinyDb</b>	<b>Cougar</b>	<b>CATMOS</b>
<b>In-network processing</b>	√	√	√
<b>Query optimization at BS</b>	×	×	√
<b>Platform used</b>	TinyOS	Query-optimizer for energy-efficient query routes	Java

### 4.3 Summary

The testing of the proposed framework ‘CATMOS’ was carried out on test-bed ‘Solarium’ along with the motes. Use of low power consuming compilation techniques have been proposed at node level and query merging at base station for optimization of data transmission in order to decrease the energy consumption for serving multiple applications by a WSN. The main objective of the study was to assess the achievement in energy savings and the objective was successfully achieved.

## Chapter-5

### Conclusion and Future Scope

The objectives of the research work have been:

- I. Review of the literature on Query processing and WSN.
- II. To propose an efficient framework for Query Processing in WSN environment.
- III. To design and develop the proposed framework.
- IV. Verify and validate the proposed framework. For e.g. by using synthetic data set.

A lot of research literature on WSN regarding history, evolution and ongoing research to overcome the inherent limitations of WSN was reviewed. On the basis of the literature reviewed it has been construed that query based techniques are the most popular and relevant means of information generation from WSN data. It has also been established from the literature that optimization of communication load is one of the most effective means of achieving energy efficiency. The research work has been formulated on these two premises. Therefore, a framework for efficient query processing in WSN has been proposed.

Most of the researchers have adopted topological and routing innovations in their work, these works have been able to achieve the desired objective but the suitability of the work has been limited to certain specific applications only. However, as the WSN today find use in diverse applications such as: ambient condition monitoring, forest fire detection, health care monitoring, battlefields intrusion or any application in any inaccessible

location etc., proposing a generalized approach has been the main motive. The proposed approach is applicable to all types of applications, irrespective of size, topography or protocols used.

In addition to this, it has been taken care of that algorithms or processing proposed do not use scarce resources of sensor nodes in compression. Some of the routing protocol based or agent based approaches introduces overheads or processing needs which negate the objective of reduction in energy consumption as processing consumes energy which is more than the saving in data reduction.

The proposed framework:“Compression At input with Multi query Optimization at Sink” (CATMOS) has been built upon query type data extraction technique which has made a two pronged attack on data redundancy i.e., In-network optimization processing at node level and BS optimization processing at BS for efficient Query Processing. In the work, element of Network management has been employed by using In-network compression.

In ‘In-network aggregation’ the data has been compressed at the node level. The main objective has been to reduce the data size by eliminating spatio-temporal redundancies. Data fidelity has been given importance therefore quantitative reduction in data was carried out without any compromise in quality of data. Reduced data to be communicated resulted into less consumption of node energy. Three compression algorithms i.e., Huffman, LZW and Deflate compression algorithms were examined on a Sun Small Programmable Object Technology (SunSPOT) network. Possible data reduction through each algorithm has been tested on a simulator ‘Solarium’.

For BS optimization, the data available at BS in response to already running queries within the network has been scrutinized to check whether the reply to newly injected queries could be generated from the data before floating the query into the network. If at all query has to be sent into the network it has been merged with the running queries and a new query has been rewritten. A novel algorithm for rewriting queries, based on the concept of Query Merging was proposed. In Query merging, new queries were rewritten along with the already running queries within the network. Only those mergers

were chosen for action, which provided actual Gains as verified through the algorithm. The merged queries were given the name 'Synthetic Queries'. The proposed algorithm ensured that the chosen synthetic queries achieved all the desired objectives. In addition to this, query language syntax modification by using single syllable static variable has been employed to achieve additional Gains.

Simulator results of compression and query optimization were encouraging. The application of this optimization resulted in a modest 10% reduction in the power consumption of the application. This much Gain has been possible through these compression algorithms which were chosen because of their low processing energy consumption.

In the future as the technology progresses the application areas of WSN will become broader than the existing scenario. Their availability to the masses will become more than right now. The applications are going to be more and more diverse with increased heterogeneity of sensors and network platforms. Generalized approaches are the best possible tools to exploit the potential of this wonderful technology. Concept of smart cities to improve quality of life using IT has acceptance world over. It is conceived on the idea of Internet of Things (IoT) such as vehicles, electronic appliances, buildings etc. which will be having wireless sensors to monitor various physical parameters. Communication among such heterogeneous objects and platforms will require standardization or approaches which are independent of nature of device or platforms.

Generalized approach used in the proposed framework maybe developed by future researchers to create more efficient applications. With time knowledge base and expertise in developing application specific energy efficient programs will increase. The proposed framework can be used along with other application specific tools to achieve further gains in the metrics of significance.

Future researchers have many obstacles to overcome to make the use of WSN to their maximum potential for the benefit of mankind. Reduction in energy and possibility of deployment in any unimaginable location should be the main focus of research.

## References

- [1] M.Weiser, “The Computer for the Twenty-first century”, *Scientific American*, Sep. 1991, pp. 94-100.
- [2] National Research Council. *Embedded, everywhere: A Research Agenda for Networked Systems of Embedded Computers*. National Academy Press, 2001.
- [3] D. Estrin, D. Culler, K. Pister, and G. Sukhatme, ”Connecting the Physical World with Pervasive Networks”,*IEEE Pervasive Computing*, , Jan 2002, pp. 59-69.
- [4] John Henkel,” Software Architecture Design of Wireless Sensor Networks”, *Basic Architecture of Wireless Sensor Network*.
- [5] Blumenthal, Jan; Handy, Matthias; Golatowski, Frank; Hasse, Marc; Timmermann, Dirk.*Wireless Sensor Networks – New Challenges in Software Engineering*. Emerging Technologies and Factory Automation,2003.Proceedings ETFA '03. IEEE Conference, Volume: 1, Sept. 2003, pp. 16-19.
- [6] Hill, Jason; Szewczyk, Robert; Woo, Alec; Hollar, Seth; Culler, David; Pister; Kristofer. *System Architecture Directions for Networked Sensors*,ASPLOS 2000.
- [7] Halawani, S., Khan, A., *Sensors Lifetime Enhancement Techniques in Wireless Sensor Networks - A Survey*.Journal of Computing, vol. 2, issue 5, May 2010.
- [8] Koubaa, A., Alves, M., Tovar , E., *Lower Protocol Layers for Wireless Sensor Networks: A Survey*. IPPHURRAYTechnical Report, HURRAY-TR-051101. 2005.
- [9] <http://rp-www.cs.usyd.edu.au/~wsn/sunspots.shtml>
- [10] Satyanarayanan, M.,” Pervasive computing: vision and challenges”, *Personal Communications*, IEEE, Volume:8 Issue:4 , Aug 2001.
- [11] Sarma, Kandarpa Kumar and AbhijitMitra. "Estimation of MIMO Wireless Channels Using Artificial Neural Networks." *Cross-Disciplinary Applications of Artificial Intelligence and Pattern Recognition: Advancing Technologies*. IGI Global, 2012, Web. 26 Sep. 2015, pp. 509-543.
- [12] Atul M. Gonsai, Bajaja, Sunil L “Ad hoc Multi-hop Wireless Networks: A Survey”, *International Journal of Computer Application* Issue 4, Volume 2 (March-April 2014).

- [13] JianPing Wang, Eseosa S. Osagie, ParimalaThulasiraman and Ruppa K. Thulasiram, "HOPNET: A Hybrid ant colony OPTimization routing algorithm for Mobile ad hoc NETWORK", Elsevier, Ad Hoc Networks, Volume 7, Issue 4, June 2009.
- [14] Menik Tissera, Robin Doss , Gang Li, Lynn M. Batten," Novel Approach for Information Discovery inAutonomous Wireless Sensor Networks", Future Network Systems and Security, Volume 523 of the series Communications in Computer and Information Science pp. 47-60.
- [15] Raymond S. Wagner, "Standards-Based Wireless Sensor Networking Protocols for Spaceflight Applications",2010 IEEE Aerospace Conference Applications and Architectures for Wireless Sensor Networks Session, March 11, 2010.
- [16] Gao, J., Guibas, L.J., Hershberger, J. and Zhang, L.: Fractionallycascaded information in a sensor network. In *Proc. of IPSN*, 2004.
- [17] I.F. Akyildiz and W. Su and Y. Sankarasubramaniam and E. Cayirci, Wireless Sensor Networks: A Survey *IEEE Computer*, vol. 38, no. 4, Mar. 2002,pp. 393-422.
- [18] <https://wirelessmeshsensornetworks.wordpress.com/>
- [19] David Gascón, Marcos Yarza," Wireless Sensor Networks to Control Radiation Levels", Published in: Articles, Environment, Makers, Smart Metering April 19th, 2011.
- [20] Alberto Bielsa, Mark Boyd," Wireless Sensor Networks to monitor food sustainability",Published in: Articles, Environment, February 28th, 2012.
- [21] E. Jovanov, et al., "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation," *Journal of Neuro Engineering and Rehabilitation*, 2005, pp. 2-6.
- [22] Martha Baer," The Ultimate on-the-fly Network", On Newsstands Now, Issue 11.12 | December 2003.
- [23] Gaura, E. and Daniel, T. (2008) Information extraction from large-scale WSNs: approaches and research issues: approaches and research issues: part I: overview and agent based approaches. *Sensors & Transducers*, volume 94 (7),pp.15-33.

- [24] Daniel, T. and Gaura, E. (2008) Information extraction from large-scale WSNs: approaches and research issues: approaches and research issues: part II: query-based and macro programming approaches. *Sensors & Transducers*, volume 94 (7), pp. 34-56.
- [25] Daniel, T. and Gaura, E. (2008) Information extraction from large-scale WSNs: approaches and research issues: approaches and research issues: part III: towards a hybrid approach. *Sensors & Transducers*, volume 94 (7), pp. 57-82.
- [26] H. Qi, Xiaoling Wang, and K. Chakrabarty, Multisensor data fusion in distributed sensor networks using mobile agents, In *5th International Conference on Information Fusion*, Annapolis, MD., 2001.
- [27] C. L. Fok, G. C. Roman, and C. Lu, Mobile agent middleware for sensor networks: an application case study, In *Proceedings of Fourth International Symposium on Information Processing in Sensor Networks* (IEEE Cat. No. 05EX1086), Los Angeles, CA, USA, 2005, pp. 382–387.
- [28] H. Qi, S. S. Iyengar, and K. Chakrabarty, Distributed multi-resolution data integration using mobile agents, In *Proceedings of IEEE Aerospace Conference 2001* (Cat. No. 01TH8542), Big Sky, MT, USA, 2001, pp. 3–1133
- [29] Robert Wesson, Frederick Hayes-Roth, John W. Burge, Cathleen Stasz, and Carl A. Sunshine, Network structures for distributed situation assessment, *IEEE Transactions on Systems, Man and Cybernetics*, SMC-11, 1, 1981, pp. 5–23.
- [30] Hairong Qi, Yingyue Xu, and Xiaoling Wang, Mobile-agent-based collaborative signal and information processing in sensor networks, In *Proceedings of the IEEE*, 91, 8, 2003, pp. 1172–1183.
- [31] Bhupinder Singh, Seema Bawa, “Grid Resource Discovery using Natural Mobile Agents”, *IJCNS, International Journal of Computer Science & Network Security*, Vol. 8, No. 11, November 2008, pp. 411 - 415.
- [32] C. Min, K. Taekyoung, C. Yanghee, Data dissemination based on mobile agent in wireless sensor networks, In *Proceedings of Conference on Local Computer Networks, LCN*, Sydney, Australia, 2005, pp. 527–528.

- [33] L. Szumel, J. LeBrun, J. D. Owens, Towards a mobile agent framework for sensor networks, In *Proceedings of 2nd IEEE Workshop on Embedded Networked Sensors, EmNetS-II*, Sydney, Australia, 2005, pp. 79–87.
- [34] Mate, Mate: Building application-specific sensor network language runtimes, 2003, <http://www.cs.berkeley.edu/~pal/research/mate.html>
- [35] Chien-Liang Fok, Gruia-Catalin Roman, and Chenyang Lu, Rapid development and flexible deployment of adaptive wireless sensor network applications, In *Proceedings of International Conference on Distributed Computing Systems*, Columbus, OH, United States, 2005, pp. 653–662.
- [36] Liu Ting and M. Martonosi, Impala: a middleware system for managing autonomic parallel sensor systems, *SIGPLAN Notices*, 38, 10, 2003, pp. 107–118.
- [37] F. Kamin Whitehouse, "A Comprehensive Approach to Macroprogramming", CENS Technical Seminar Series, Boelter Hall 4760, January 9, 2009.
- [38] Ramakrishna Gummadi, Omprakash Gnawali, Ramesh Govindan, "Macroprogramming Wireless Sensor Networks Using *Kairos*", *Distributed Computing in Sensor Systems* Volume 3560 of the series *Lecture Notes in Computer Science*, pp. 126-140
- [39] A.Hopper. The Clarendon Paterson Lecture: Sentient Computing. *Philosophical Transactions of the Royal Society of London*, 358(1773), August 1999, pp. 2349-2358.
- [40] Yao, Y., Gehrke, J: "The COUGAR approach to in-network query processing in sensor networks", *SIGMOD Rec.* 31(3) (200) 9-18.
- [41] Shen Chien-Chung, C. Srisathapornphat and C. Jaikaeo, "Sensor Information networking architecture and applications", *IEEE Personal Communication*, 8(4), 2001, pp 52-59.
- [42] L. Shuoqi, S.H. Jon, J.A. Stankovic, "Event detection services using data service middleware in distributed sensor networks", In *Proceedings of Information Processing in Sensor Networks. 2<sup>nd</sup> Int'l Conf.*
- [43] J. Gehrke and S. Madden, "Query processing in sensor networks," *IEEE Pervasive Computing*, Jan-Mar. 2004, pp. 46-55.

- [44] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TinyDB: An acquisitional query processing system for sensor networks," *ACM Trans. On Database Systems*, Vol. 30, Issue 1, Mar.2005, pp.122-173.
- [45] P. Bonnet, J. Gehrke, P. Seshadri, "Querying the physical world", *IEEE Personal Communications*, 7, 5, 2000, pp. 10–15.
- [46] ArnabChakrabarti, AshutoshSabharwal, and BehnaamAazhang, "Using Predictable Observer Mobility for Power Efficient Design of Sensor Networks", (2002)
- [47] IoannisChatzigiannakis, AthanasiosKinalis , Sotiris Nikolettseas "Efficient data propagation strategies in wireless sensor networks using a single mobile sink", Elsevier Computer Communication , (2008).
- [48] Yan Wu , Fahmy, Sonia , Shroff, N.B.," Optimal Sleep/Wake Scheduling for Time-Synchronized Sensor Networks With QoS Guarantees ",Published in Networking, IEEE/ACM Transactions on (Volume:17 , Issue: 5 ), 30 June 2009,pp. 1508 – 1521.
- [49] Feng Wang and Jiangchuan Liu," Duty-Cycle-Aware Broadcast in Wireless Sensor Networks", Published in INFOCOM 2009, IEEE, Rio de Janeiro, 19-25 April 2009, pp 468 – 476.
- [50] Yan Wu, Fahmy, S.; Shroff, N.B.," Energy Efficient Sleep/Wake Scheduling for Multi-Hop Sensor Networks: Non-Convexity and Approximation Algorithm", Published in INFOCOM 2007, 26<sup>th</sup> IEEE International Conference on Computer Communications. IEEE, Anchorage, AK, 6-12 May 2007, pp. 1568 - 1576
- [51] Babar Nazir, HalabiHasbullahandSajjad A Madani, "Sleep/wake scheduling scheme for minimizing end-to-end delay in multi-hop wireless sensor networks", In *EURASIP Journal on Wireless Communications and Networking* (EURASIP JWCN, 2011, 2011:92 doi:10.1186/1687-1499-2011-92
- [52] Issa Khalil, SaurabhBagchi, Ness B. Shroff,"SLAM : Sleep-Wake Aware Local Monitoring in Sensor Networks, "In 37<sup>th</sup> Annual IEEE/IFIP International

- Conference on Dependable Systems and Networks, 2007. DSN '07, Edinburgh, 25-28 June 2007, 565 – 574.
- [53] Z. Liu, B. Huffaker, M. Fomenkov, K. Claffy, and N. Brownlee, "Two Days in the Life of the DNS Anycast Root Servers", Passive and Active Network Measurement Workshop (PAM), vol. 4427, Apr 2007, pp. 125--134.
- [54] Dina Katabi, Wroclawski, J.," A framework for scalable global IP-anycast (GIA)", In Proceedings of the conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, ACM NY, USA 2000, pp. 3-15.
- [55] Tapaswini Dash, Bharati Mishra," A Hybrid Approach of Using Anycast Addressing With Zone Routing Protocol", In IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 2, July 2012.
- [56] Hussein A. Alzoubi, SeungjoonLee,MichaelRabinovich, Oliver Spatscheck, Jacobus Van der Merwe," Anycast CDNs Revisited", www 2008/ Refereed Track, April 21-25, 2008, Beijing, China.
- [57] ArmelEsnault , Nicolas Le Sommer , FrédéricGuidec ," An Anycast Communication Model for Data Offloading in Intermittently-Connected Hybrid Networks , In Procedia Computer Science, Volume 56, 2015, pp. 59–66.
- [58] M. Cardei and D.Z. Du, "Improving wireless sensor network lifetime through power aware organization", ACM Wireless Networks , vol. 11, no. 3., 2005, pp. 333–340.
- [59] Y. Zou and K. Chakrabarty, "A distributed coverage- and connectivity-centric technique for selecting active nodes in wireless sensor networks," IEEE Trans. Computers , vol. 54, no. 8, 2005, pp. 978–991.
- [60] Q. Zhao and M. Gurusamy, "Lifetime maximization for connected target coverage in wireless sensor networks", IEEE/ACM Trans. Net-working , vol. 16, no. 6, 2008, pp. 1378–1391.
- [61] N. Heo and P.K. Varshney, "Energy-efficient deployment of intelligent mobile sensor networks", IEEE Trans. Systems, Man and Cybernetics – Part A: Systems and Humans , vol. 35, no. 1,2005, pp. 78–92.

- [62] Y.C. Wang, C.C. Hu, and Y.C. Tseng, "Efficient placement and dispatch of sensors in a wireless sensor network", *IEEE Trans. Mobile Computing*, vol. 7, no. 2, 2008, pp. 262–274.
- [63] Rice, J. A., K. A. Mechitov, S. H. Sim, B. F. Spencer, and G. A. Agha," Enabling frame-work for structural health monitoring using smart sensors",*Structural Control and Health Monitoring*. doi:10.1002/stc.386. (2010)
- [64] Lynch, J. P., Sundararajan, A., Law, K.H., Kiremidjian, A.S., and Carryer, E.," Embedding damage detection algorithms in a wireless sensing unit for operational power efficiency", *Smart Materials and Structures*, 13, no. 4, pp. 800-810. doi:10.1088/0964-1726/13/4/018. (2004)
- [65] C Alippi, G Anastasi, M Di Francesco, M Roveri," Energy management in wireless sensor networks with energy-hungry sensors" ,*Instrumentation & Measurement Magazine, IEEE* 12 (2), Apr 2009, pp.16-23,.
- [66] Sangeeta Rani and GeetaSikka. Article: Recent Techniques of Clustering of Time Series Data: A Survey. *International Journal of Computer Applications* 52(15), August 2012, pp. 1-9.
- [67] Sally Ida McClean , LalitGarg, PriyankaCharasia, Bryan W. Setney, Chris D.Nugent, "Using model-based clustering to discretise duration information for activity recognition", In *Proceedings of the 24<sup>th</sup> IEEE International Symposium on Computer-Based Medical Systems*, 27-30 June, 2011, Bristol, United Kingdom.
- [68] S Seo, J Kang, D Lee, KH Ryu," Multivariate stream data classification using simple text classifiers", *DEXA 2006, LNCS 4080*, 2006, pp. 420 – 429.© Springer-Verlag Berlin Heidelberg 2006.
- [69] M. Nelson and J. L. Gailly,*The data compression book*, MIS Press, 1996.
- [70] D.S. Taubman and M.W. Marcellin,*JPEG2000: fundamentals, standards and practice* , Kluwer Academic Publishers, 2002
- [71] Chung-E Wang," Dynamic LZW for Compressing Large Files ", athena.ecs.csus.edu
- [72] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang. MACAW: A media access protocol for wireless LANs. In *Proceedings of ACM SIGCOMM*, 1994.

- [73] Rongxing Lu, Xiaodong Lin, Haojin Zhu, Xiaohui Liang," BECAN: A Bandwidth-Efficient Cooperative Authentication Scheme for Filtering Injected False Data in Wireless Sensor Networks", *Parallel and Distributed Syst*, Volume:23 Issue:1.
- [74] F. F.E. White, *Data Fusion Lexicon*, Joint Directors of Laboratories, Technical Panel for C3, Data Fusion Sub-Panel, Naval Ocean Systems Center, San Diego,1991.
- [75] L.A. Klein, *Sensor and Data Fusion Concepts and Applications*, second ed., Society of Photo-optical Instrumentation Engineers (SPIE), Bellingham, WA, 1999.
- [76] B.V. Dasarathy, *Decision Fusion*, IEEE Computer Society Press, Los Alamitos, CA, 1994.
- [77] I.R. Goodman, R.P.S. Mahler, H.T. Nguyen, *Mathematics of Data Fusion*, Kluwer Academic Publishers, Norwell, MA, 1997.
- [78] Heinzelman, W., Chandrakasan, A., and Balakrishnan, H., "Energy-Efficient Communication Protocols for Wireless Microsensor Networks", *Proceedings of the 33rd Hawaiian International Conference on Systems Science (HICSS)*, January 2000.
- [79] Samuel Madden, Michael J. Franklin, and Joseph M. Hellerstein Wei Hong," TAG: aTiny Aggregation Service for Ad-Hoc Sensor Networks", *Appearing in 5<sup>th</sup>Annual Symposium on Operating Systems Design and Implementation (OSDI)*. December, 2002
- [80] K.Sohrabi, J.Gao, Y.Ailawadhi and G.J.Pottie," Protocols organization of Wireless Sensor Network", in *Proceedings of IEEE Personal Communications*, 7(5), Oct. 2000, pp. 16-27,.
- [81] C.Intanagonwivat, R.Govondan and D.Estrin," Directed Diffusio: A Scalable and Robust Communication paradigm for sensor networks", *In Proceedings of the ACM/IEEE Int'l Conf. on Mobile Computing and Network (Mobicom)*, Boston, Aug 2000.

- [82] P.Bonnet, J.Gehrke and P.Sheshadri, "Towards sensor database systems, mobile data management", 2<sup>nd</sup> International Conference, MDM 2001, Proceedings (LNCS vol. 1, 1987), Hongkong, China, Springer-Verlay, 2001, pp 3-14.
- [83] Sadagopan, N; Krishnamachari, B.; Helmy,A., "The ACQUIRE Mechanism for Efficient Querying in Sensor Networks", In Proceedings of First IEEE Int'l Workshop on Sensor Network Protocols and Applications (SNPA), Anchorage, AK, May 2003; pp 49-155.
- [84] A.Coman, M.A.nascimento, J.Sander, "A framework for Spatio-temporal query processing over wireless sensor network", In DMSN '04, *Proceeding of the Int'l workshop on Data Management for sensor network*.
- [85] M.Demirbas, X.Lu, P.singla,"An In-Network Querying Framework for Wireless Sensor Networks", IEEE Transactions on Parallel and Distributed Systems.
- [86] Y.Chen, M.Xu, H.Wang, H.V.Leong, J.Cao, K.C.C.Chan and A.T.Schan, "An Energy Efficient Framework for Multi-rate Query in Wireless Sensor Networks" *EURASIP Journal of Wireless Communications and Networking*, May 2007.
- [87] A.Tavakoli, "Wringer: A debugging and Monitoring Framework for Wireless Sensor Networks", *Proceedings of the ACM SenSys Doctoral Colloquium*, 2007.
- [88] P.Kumar, A.J. Choudhury, M.Sain, S.Lee and H.Lee,"RUASN: A Robust User Authentication Framework for Wireless Sensor Networks",*Sensors2011*,11(5), 4 May 2011.
- [89] H.Jiang, S.Jin, C.Way, "Prediction or not? An Energy Efficient Framework for Clustering Based Data Collection in Wireless Sensor Networks", vol. 22, no. 6, Jun 2011, pp 1064-1071.
- [90] S.Pervin, J.Kamruzzaman, G.Karmakar, A.K.M.Azad," Hybrid In-Network Query Processing Framework for Wireless Sensor Networks", Communication (ICC), 2011 IEEE Int'l Conference.
- [91] A. Qandour (2012), "Application framework for Wireless Sensor networks", Retrieved from <http://ro.cui.au/theses/472>.
- [92] L. Sarakis, T.Zahariadis, H.Leligou, M.Dohler, "A Framework for Service Provisioning in virtual sensor networks", *EURASIP Journal on Wireless Communications and Networking* 2012.

- [93] IEEE Computer Society LAN MAN Standards Committee. Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. In IEEE Std. 802.11,1999.
- [94] S.Singh and C. Raghavendra. PAMAS: Power aware multi-access protocol with signaling for ad hoc networks. ACM SIGCOMM Computer Communication Review, 28(3), July 1998, pp. 5–26.
- [95] J.E.B.Maia, A.Brayner and F.Lodrigues, "A framework for processing complex queries in wireless sensor networks", Newsletter, ACM SIGAPP Applied Computing Review, Vol. 13, Issue 2, Jun 2013, pp.30-41.
- [96] Woehrle M., Plessl C., Lun R., Bentel J., Thiele L.," EvAnT: Analysis and Checking of event traces for Wireless Sensor Networks", 2008 IEEE Int'l Conf. on Sensor Networks, Ubiquitous and Trustworthy Computing.
- [97] O.Chipra, C.Lu and J.A. Stankovich, " Dynamic conflict Frequency scheduling for Wireless Sensor Networks", In ICNP, 2006.
- [98] C. Min, K. Taekyoung, C. Yanghee, Data dissemination based on mobile agent in wireless sensor networks, In *Proceedings of Conference on Local Computer Networks, LCN*, Sydney, Australia, 2005, pp. 527–528.
- [99] Java book Programming with Java by E Balagurusamy
- [100] Madden, S. and Franklin,M.J.," Fjording the stream: An architecture for queries over streaming sensor data", In ICDE, San Jose, CA, February 2002.
- [101] <https://www.quora.com/What-is-wireless-sensor-network-WSN-technology>
- [102] <http://ccna-cme.blogspot.in/2012/06/day-4-models.html>
- [103] WIRELESS SENSOR NETWORKS A Networking Perspective Book, Edited by Jun Zheng, Abbas Jamalipour.
- [104] <http://rtcmagazine.com/articles/view/102871>
- [105] <http://www.elprocus.com/introduction-to-wireless-sensor-networks-types-and-applications/>
- [106] <http://www.temperaturealert.com/Wireless-Temperature-Store/Temperaturelert-Dry-Media-Temperature-Buffer-Vial-Medium-P80.aspx>
- [107] <file:///C:/DOCUME~1/Pc/LOCALS~1/Temp/SunSPOT-Programmers-Manual.htm>