ENERGY-EFFICIENT DATA TRANSMISSION WITH CLUSTERING AND COMPRESSIVE SENSING IN WIRELESS SENSOR NETWORKS

MUKIL ALAGIRISAMY

FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

ENERGY-EFFICIENT DATA TRANSMISSION WITH CLUSTERING AND COMPRESSIVE SENSING IN WIRELESS SENSOR NETWORKS

MUKIL ALAGIRISAMY

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **MUKIL ALAGIRISAMY**

Matric No: **KHA130024**

Name of Degree: **DOCTOR OF PHILOSOPHY**

Title of Thesis: **ENERGY-EFFICIENT DATA TRANSMISSION WITH CLUSTERING AND COMPRESIVE SENSING IN WIRELESS SENSOR NETWORKS**

Field of Study: **WIRELESS SENSOR NETWORKS AND COMMUNICATION**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 1/6/2020

Subscribed and solemnly declared before,

Witness's Signature Date: 1/6/2020

Name:

ENERGY-EFFICIENT DATA TRANSMISSION WITH CLUSTERING AND COMPRESSIVE SENSING IN WIRELESS SENSOR NETWORKS

ABSTRACT

One of the most important application of wireless sensor network is environmental monitoring. The application involves lifetime of sensor nodes for longer duration associating its energy module. Wireless sensor nodes deployed in sensing field aggregate enormous amount of sensed data and transfer them to the sink. The inherent limitation of energy carried within the battery of sensor nodes fetches extreme difficulty to acquire adequate network lifetime, becoming a bottleneck in forwarding data to sink. Hence the motivation is to reduce the amount of data transfer and attain energy efficiency. This is achieved by clustering and compressive sensing techniques. First objective is to reduce the transmission burden on sensors to attain energy efficiency. The solution is achieved by unequal clustering with appropriate cluster head selection and dual sink. These two criterions minimizes the energy holes and preserves the network lifetime. Energy-Aware Unequal Clustering routing algorithm with Dual sink (EAUC-DUAL) and Energy based Cluster Head selection Unequal Clustering with Dual sink (ECH-DUAL) are proposed. EAUC –DUAL uses static and mobile sink. EAUC-DUAL suggests smaller size clusters around static sink and the entire clusters in the network transmit their data only to the nearest sink providing load balancing and minimizes the hotspot. In the extended ECH-DUAL algorithm in addition to dual sink a new cluster head selection method is proposed for unequal clustering. It focuses on balancing the burden of cluster heads by suitable selection of Tentative Cluster Head (TCH) and Final Cluster Head (FCH). Simulation results interprets the network lifetime of EAUC –DUAL is two times more than the LEACH algorithm. In the extended ECH-DUAL algorithm the network lifetime of ECH-DUAL is twice greater than the network lifetime of EAUC-DUAL.

Second objective is to reduce the transfer volume of sensed data and attaining energy efficiency. Conventional sampling results in high amount of sensed data. Hence the framework of Intelligent Neighbor-Aided Compressive Sensing (INACS) is proposed emphasizing on compressive sensing at source, data forwarding based on highest correlation to the neighbor node and exact recovery at the sink. Compressive sensing techniques and data forwarding reduces the volume of sensed data and the number of transmissions. Simulation results conclude better energy efficiency and reconstruction accuracy. The energy consumption in INACS is 0.29 times lesser than the existing protocol.The third objective focus on the reduction in observational cost and transmission cost through compressive sensing and data forwarding techniques respectively. The data forwarding should be performed considering the link capacity of nodes and available bandwidth. A framework Perceptron-based Optimal Routing (POR) and Perceptron-based Routing with Moderate Traffic Intensity (PRMTI) is proposed for data forwarding. The data forwarding process is initiated considering the network resources. POR is suggested for scarce network resources and PRMTI for abundant network resources. Simulation has been performed on energy consumption and number of transmissions. Residual energy of POR is 0.26 times higher and PRMTI is 0.14 times higher than the existing protocol. The simulation results of clustering algorithms, INACS and Perceptron framework are validated using various data analysis methods.

Keywords: Clustering, Dual Sink, Compressive Sensing, Data Forwarding, Energy Consumption and Network Lifetime.

PENGHANTARAN DATA BERKECEKAPAN TENAGA DENGAN PENGGUGUSAN DAN PENGESAN MAMPATAN DIDALAM RANGKAIAN SENSOR TANPA-WAYAR

ABSTRAK

Salah satu aplikasi rangkaian sensor tanpa-wayar yang paling penting adalah di dalam pemantauan persekitaran. Aplikasi ini melibatkan masa hayat nod sensor untuk tempoh yang lebih lama dengan mengaitkan modul tenaga di dalam nod sensor. Nod sensor tanpa-wayar digunakan ke atas data yang dikesan di dalam kawasan pengesanan yang besar dan memindahkannya ke sink. Had tenaga yang dibawa oleh bateri di dalam nod sensor menyebabkan kesukaran untuk mendapat masa hayat yang cukup, ini menyebabkan kesesakan penghantaran data ke sink. Oleh yang demikian, tujuan motivasi ini adalah mengurangkan jumlah pemindahan data dan mencapai kecekapan tenaga. Ini dapat dicapai dengan teknik penggugusan dan penderian mampatan. Objektif pertama ialah mengurangkan bebanan penghantaran ke atas sensor untuk mendapatkan kecekapan tenaga. Ini dapat diselesaikan dengan menggunakan penggugusan tak sama dengan pemilihan ketua gugusan yang sesuai dan dual sink. Kedua-dua kriteria ini mengurangkan lubang tenaga dan mengekalkan masa hayat rangkaian. Algoritma penghalaan Energy-Aware Unequal Clustering with Dual sink (EAUC-DUAL) dan Energy based Cluster Head selection Unequal Clustering with Dual sink (ECH-DUAL) telah dicadangkan. EAUC-DUAL menggunakan sink yang statik dan bergerak. EAUC-DUAL menggunakan saiz gugusan yang kecil sepanjang sink yang statik dan kesemua gugusan di dalam rangkaian menghantar data hanya kepada sink yang terdekat dengan menyediakan keseimbangan beban dan mengurangkan “titik panas”. Algoritma ECH-DUAL turut dicadangkan di dalam pemilihan kepala gugusan untuk penggugusan tidak sama. Ia tertumpu kepada keseimbangi beban kepala gugusan dengan pemilihan Tentative Cluster Head (TCH) dan Final Cluster Head (FCH). Keputusan simulasi menyatakan masa hayat rangkaian EAUC-DUAL adalah dua kali ganda lebih besar daripada algoritma LEACH.

Di dalam algoritma ECH-DUAL, masa hayat rangkaian ECH-DUAL adalah dua kali ganda lebih besar daripada masa hayat rangkaian EAUC-DUAL. Objektif kedua adalah mengurangkan jumlah pemindahan data yang telah dikesan dan mencapai kecekapan tenaga. Persampelan lazim menyebabkan jumlah yang tinggi di dalam data yang telah dikesan. Oleh yang demikian, rangka kerja Intelligent Neighbor-Aided Compressive Sensing (INACS) dicadangkan, di mana ia memberi penekanan kepada penderiaan mampatan di sumber, penghantaran data berdasarkan korelasi tertinggi kepada nod jiran dan pemulihan yang tepat di sink. Penderiaan mampat dan penghantaran data mengurangkan jumlah data yang dikesan dan bilangan penghantaran. Keputusan simulasi menyimpulkan kecekapan tenaga dan peratusan pembentukan semula adalah bagus. Penggunaan tenaga di dalam INACS adalah 0.29 kali kurang daripada protokol sedia ada. Objektif ketiga tertumpu kepada pengurangan jumlah pemantauan dan penghantaran melalui teknik penderian mampat dan teknik penghantaran data. Penghantaran data mesti dilakukan dengan mengambil kira kapasiti laluan nod dan jalur lebar tersedia. Rangka kerja Perceptron-based Optimal Routing (POR) dan Perceptron-based Routing with Moderate Traffic Intensity (PRMTI) dicadangkan untuk penghantaran data. Proses penghantaran data dijalankan dengan mengambil kira sumber-sumber rangkaian. POR dicadangkan untuk sumber-sumber rangkaian yang kurang dan PRMTI untuk sumber-sumber rangkaian yang banyak. Simulasi telah dijalankan ke atas penggunaan tenaga dan bilangan penghantaran. Tenaga residu POR adalah 0.26 lebih tinggi dan PRMTI adalah 0.14 lebih tinggi daripada protokol sedia ada. Keputusan simulasi algoritma penggugusan, INACS dan rangka kerja Perceptron adalah terbukti dengan menggunakan pelbagai teknik analisis data.

Kata kunci: Penggugusan, Dual Sink, Masa Hayat Rangkaian, Pengesan Mampatan, Penghantarab Data, Penggunaan Tenaga.

ACKNOWLEDGEMENTS

Writing this thesis is one of the important milestones on the long journey of pursuing my doctoral degree, and this journey would have been harder without the encouragement from many individuals. I would like to present my sincere acknowledgements and deep regards to those who have supported me.

First of all I would like to convey my deep sense of gratitude from the bottom of my heart to my respectable and honorable supervisor Dr.Chow Chee Onn who have been the driving force and guiding force behind the accomplishment of my thesis. I am deeply indebted to him for proposing me the cutting edge of research in wireless sensor networks and offering me the precious opportunity to accomplish my doctorate degree under his supervision. His vast knowledge in wireless sensor networks and great experience in research have assisted me tremendously throughout the accomplishment of my doctorate degree. His flexibility, continuous support and encouragement during all my troubles throughout study have paved way to submit my thesis finally. His technical and editorial suggestions are crucial to the completion of this thesis dissertation. My special words of thanks to Dr. Kamarul Ariffin Bin Noordin for his guidance and support at all times.

Furthermore, I would like to express my sincere gratitude to the management and Faculty of Engineering of Lincoln University College. I would like to thank our CEO Dr. Amiya Bhaumik, Dean of Faculty of Engineering Datuk Ismail Bin Hassan for understanding my difficulties during thesis writing and approving me the necessary study leaves. Thank you to my dear friend Ms. Rohilah Sahak for helping me in translations during thesis writing. Ms. Rohilah you have been a great pillar of strength from the day I knew you. I am very proud of you my friend. I would also like to extend my gratitude to my friends and colleagues Ms. Supraja, Mr. Kalai, Mr. Antony and Mr. Nizam for substituting my work at the Faculty of Engineering and allowing me to write my thesis peacefully without any worries. I am also greatly indebted to my brothers in India, Mr.

Balaguru, Mr. Anand and Mr. Siva for offering insightful advice throughout the accomplishment of my research.

I can never give enough thanks to my beloved parents: my dearest father Dr. Alagirisamy and my lovely mother Madam Ezhilarasi. They have been insisting me to pursue Ph.D. and they take care of my children in India to diminish my personal works, so that I can complete my doctorate degree successfully. They would be the happiest persons in the world once I finish my degree. A simple thanks will not equalize their sacrifices. I am really gifted to be their daughter and I pray the same continues in every birth of mine. To my elder sister Dr. Chinthanai, thank you for being a companion not only during our childhoods but over the lifetime. She had been a role model to accomplish my thesis without over stressing myself.

My heartfelt gratitude is to my soulmate, best friend, colleague and life partner Mr.Balasubramaniyan Seetharaman who is always a sympathetic listener, motivates, helps me in house hold works in spite of his busy schedules and keeps me out from all issues so that I complete my doctorate degree. Without him, I can never expect to achieve what seems impossible to begin with. Thank you for your endless care and affection.

Last but not least thanks to my lovable nephews Mr. Agaran, Mr. Anban and Mr. Adhiban for making me happy and cool during all the phone calls from India throughout my research work. My sincere apologies to my brilliant, naughty, cute, dear children Mr. Niruban and Ms. Amizhthini for leaving you in India due to my tight work schedules. Now I am very much happy and excited to join you. I would like to end saying never ever give up your goals. Overcome all the obstacles on your way to reach the goals. The happiness you gain after achieving your goal is the finest and priceless experience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Symbols and Abbreviations

E_{Tx}	:	Transmitted Energy
d	:	Distance
E_{Rx}	:	Received Energy
E_{elec}	:	Dissipated energy in a circuit
ε_{fs}	:	Free space fading
ε_{mp}	:	Multipath fading
R_t	:	Transmission range
R_s	:	Sensing range
C_i	:	Communication data
D	:	Amount of instructions in transfer of data to the sink
B_w	:	Bandwidth
α	:	Constant value of power consumption for communication
β	:	Constant value of power consumption for instruction transfer
l_0	:	l_0 normalization
l_1	:	l_1 normalization
m	:	Number of measurements
s	:	Non-zero coefficients
N_l	:	Length of signal
E_{sp}	:	Energy required for processing
ε_{mr}	:	Memory reading
ε_{mw}	:	Memory writing
N_o	:	Original signal acquired
M_o	:	Compressed signal
S	:	Primary cluster heads

L	:	Links
E_{DA}	:	Energy consumed in aggregating data
d_{min}	:	Minimum distance to sink
d_{max}	:	Maximum distance to sink
R_r	:	Competition range
R_o	:	Maximum competition range
N_s	:	Number of sensor nodes
N_{FD}	:	Number of forwarded packets
N_{RD}	:	Number of received packets
d(n)	:	Degree of node
N_s	:	Number of sensor nodes
φ_{mp}	:	Transmit Amplifier
df	:	Degree of freedom
Exp(B)	:	Hazard Function
SE	:	Standard Error
CI	:	Confidence Interval
N_{sa}	:	Number of samples
W_c	:	Channel Bandwidth
X	:	Unknown signal
$\bar{y}$	:	Observation matrix
ϕ_s	:	Sensing matrix
R	:	Correlation measure
A_i	:	Spatial coordinates
B_i	:	Temporal coordinates
$\overline{A}$	:	Mean of Spatial coordinates

$\bar{B}$	:	Mean of Temporal coordinates
I	:	Identity matrix
ψ_s	:	Sparsity basis
S_i	:	Packets generated at a source sensor
S_d	:	Packets generated at a destination sensor
w_{1R}	:	Weight value of node 1
w_{2R}	:	Weight value of node 2
EAUC	:	Energy-Aware Unequal Clustering
ECH	:	Energy-based Cluster Head selection
INACS	:	Intelligent Neighbor-Aided Compressive Sensing
POR	:	Perceptron-based Optimal Routing
PRMTI	:	Perceptron-based Routing with Moderate Traffic Intensity
SNR	:	Signal-to-Noise Ratio
PCH	:	Primary cluster head
SCH	:	Secondary cluster head
QoS	:	Quality of Service
SQL	:	Structured Query Language
ARSCs	:	Adjustable range set covers
DiLCO	:	Distributed Lifetime Coverage Optimization
CAG	:	Clustered Aggregation technique
DSC	:	Distributed source coding
RDRs	:	Regular Data Reports
UENs	:	Urgent Event Notifications
MLE	:	Maximum Likelihood Estimate
OSI	:	Open System Interconnect

BMAC	:	Bidirectional Medium Access Control
RIP	:	Restricted Isometry Property
JSM-1	:	Joint Sparsity Model-1
JSM-2	:	Joint Sparsity Model-2
MMV	:	Multiple Measurement Vector
CR	:	Compression Ratio
DCS	:	Distributed Compressive Sensing
ADC	:	Analog-to-Digital Converter
DAC	:	Digital-to-Analog converter
DSP	:	Digital Signal Processing
CDA	:	Compressive Data Aggregation
AR	:	Auto Regression
DACS	:	Data Acquisition and Compressive Sensing
CIDPS	:	Compressive In-Network Data Processing
ODECS	:	On-Demand Explosion-based Compressive Sensing
CSPR	:	Compressive Sensing-based Path Reconstruction
CSPLM	:	Compressive Sensing based on Packet Loss Matching
CCS	:	Clustered Compressive Sensing
BDM	:	Block Diagonal Measurement
ECOCS	:	Energy Consumption and Optimized Compressive Sensing
HDACS	:	Hierarchical Data Aggregation and Compressive Sensing
FTCS	:	Forwarding Tree Construction and Link Scheduling
TADA	:	Topology-Aware Data Aggregation
VSDA	:	Vandermonde Matrix-Based Scalable Data Aggregation
CStoragePB	:	Compressive Data Storage Probabilistic Broadcasting

CStorageAB	:	Compressive Data Storage Alternative Branching
EBT	:	Energy-Based Timer
TV	:	Trust Value
AWGN	:	Additive White Gaussian Noise
EBDG	:	Energy-Balanced Data Gathering
DCMDC	:	Dynamic Clustering Mobile Data Collector
MTE	:	Minimum Transmission Energy
MS	:	Mobile Sink
CM	:	Cluster Member
SS	:	Static Sink
NPCH	:	Neighbor Primary Cluster Head
FCH	:	Final Cluster Head
TCH	:	Tentative Cluster Head
RE	:	Residual Energy
ARE	:	Average Residual Energy
TEC	:	Total Energy Consumption
EEUC	:	Energy Efficient Unequal Clustering
ECH_DUAL_A	:	Energy-Based Cluster Head Selection Alive Nodes
ECH_DUAL_AE	:	Energy-Based Cluster Head Selection Average Energy
EEUC_DUAL_A	:	Energy Efficient Unequal Clustering Alive Nodes
EEUC_DUAL_AE	:	Energy Efficient Unequal Clustering Average Energy
CWS	:	Compressive Wireless Sensing
IHT	:	Iterative Hard Thresholding
DCT	:	Discrete Cosine Transform
DFT	:	Discrete Fourier Transform

MECDA	:	Minimum Energy Compressed Data Aggregation
NCS	:	Non-Uniform Compressive Sensing
ANCS	:	Adaptive Non-Uniform Compressive Sampling
CDG	:	Compressive Data Gathering
PCA	:	Principal Component Analysis
STCDG	:	Spatial Temporal Compressive Data Collection Scheme
HCS	:	Hybrid Compressive Sensing
TCCS	:	Topologically Conjugate Chaotic System
SRM	:	Structurally Randomized Matrix
NACS	:	Neighbor-Aided Compressive Sensing
DDCD	:	Data Density Correlation Degree
WSTCS	:	Weighted Spatial-Temporal Compressive Sensing
SCORE	:	Sensing Compression And Recovery Through Online Estimation
DASS	:	Distributed Adaptive Compressive Sensing
CACS	:	Cost-Aware Compressive Sensing
RCS	:	Regularized Column Sum
BCS	:	Bayesian Compressive Sensing
SRI	:	Sampling Rate Indicator
SPSS	:	Statistical Package For Social Sciences
ARIMA	:	Auto Regressive Integrated Moving Average
ACF	:	Auto Correlation Function
PACF	:	Partial Auto Correlation Function
RMSE	:	Root Mean Square Error
BIC	:	Bayesian Information Criteria

MAE	:	Mean Absolute Error
MAPE	:	Mean Absolute Percentage Error
MaxAPE	:	Maximum Absolute Percentage Error
MaxAE	:	Maximum Absolute Error
SE	:	Standard Error
MAPE	:	Mean Absolute Percentage Error
EC_INACS	:	Energy Consumption For Intelligent Neighbor Aided Compressive Sensing
EC_CWS	:	Energy Consumption For Compressive Wireless Sensing
DWT	:	Discrete Wavelet Transform
POR	:	Perceptron-Based Optimal Routing
PRMTI	:	Perceptron-Based Routing With Moderate Traffic Intensity
DCE	:	Distributed Compressed Estimate
MSE	:	Mean Squared Error
TCS	:	Tensor Based Compressive Sensing
LACD-EDC	:	Layered Adaptive Compressive Design For Efficient Data Gathering
RWCS	:	Random Walk Based Compressive Sensing
LESS	:	Link Estimation With Sparse Sampling
LWCDA	:	Light Weight Compressed Data Aggregation
SRMF	:	Sparsity-Regularized Matrix Formulation
EVD	:	Eigen Value Decomposition
SBL	:	Sparse Bayesian Learning
MMV	:	Multiple Measurement Vectors
MINLP	:	Mixed Integer Nonlinear Programming
NLP	:	Nonlinear Programming

CNCDS	:	Compressed Network Coding Based Distributed-Data Storage
CB CS	:	Covariogram-Based Compressive Sensing
Ideg	:	Integrated Data and Energy Gathering
POR_RE	:	Perceptron-Based Optimal Routing with Residual Energy
POR_EC	:	Perceptron-Based Optimal Routing with Energy Consumption
PRMTI_RE	:	Perceptron-Based Routing with Moderate Traffic Intensity with Residual Energy
PRMTI_EC	:	Perceptron-Based Routing with Moderate Traffic Intensity with Energy Consumption
ECTRA	:	Energy Conserving Transmission Radius Adaptive

CHAPTER 1: INTRODUCTION

1.1 Overview

Wireless sensor nodes are typically embedded in an environment and it provides information about physical phenomena rather than human interaction with the environment. The wireless sensor nodes consist of a sensing unit and communication unit. The sensing unit measures the impact of physical phenomena and communication unit transfers the data to actuator or sink. Sink are entities procuring the sensed data from sensor nodes based on the physical phenomena. Sensors are being positioned in the environment and they report the status of information to sink either periodically or at the occurrence of event. Wireless sensors are smaller in size and supply energy to meet the demands of both sensing and communication unit (Akyildiz et al., 2002). Investing more energy to meet quality of service of sensed data can result in shorter lifetime. The scope of this thesis concerns about the estimation of sensor lifetime and saving the energy consumption within the deployed sensing field (terrain).

The major issue associated with sensor node lifetime is energy consumption. The interlinked factors related to energy consumption are sensing process and communication process. The problem behind the sensing approach is, it generates a huge amount of sensor readings and has to be transferred to the sink. The sensors with resource constraints are subjected to energy drainage due to physical data acquisition and subsequent forwarding (Vuran et al., 2004). Thus the assisting solution to overcome energy drainage and transfer of sensed data can be complete using the attributes of clustering and dual sink. The clustering technique selects a cluster head to transmit data thereby reducing the flow of data from all its cluster members to the sink. The addition of mobile sink with the existing static sink reduces the hot spot problems.

The framework of clustering and dual sink leads to the reduction in transmission cost however issues with observational cost has not been considerably reduced. Transmission cost refers to the process of transfer of data from the sensors to the sink. Observational cost is the method of acquiring data from a sensing field using the desired sensing threshold. The observational cost of data acquisition can be reduced by compressive sensing techniques. The nature of compressive sensing incorporates sparsity where in signal can be represented in a miniature form without losing the quality (Aeron et al., 2006). In the scenario where the signal is not sparse the sparsity can be attained by means of certain transformations.

The compressive sensing techniques significantly reduce the local computation and sensor data volume. Thus we can say, both transmission cost and observational cost is reduced due to the compression incurred. However still losses occur during transmission due to attenuation and channel impulse response. The assisting solution should govern the status of channel impairment, forwarding at intermediate nodes, post processing accuracy for reconstruction of sparse physical data. Achieving sparse representation and transfer of data to sink considering the capacity of network resources are crucial in the compressive sensing framework.

1.2 Motivation

The motivation behind this primary work is centered on transmission cost indicating a reduction in data transmission with a suitably elected cluster head governing the metrics of time and energy efficiency. Unequal cluster sizes in the network does not allow the sensor nodes to die out faster by saving their energy. However sensors with resource constraints are still subject to time varying energy drain. To counter this problem a full operation of sinks (static and mobile sinks) has to be used with a network infrastructure to collect data. The dual sink prevents the energy holes and increases the network lifetime saving the energy consumed inside the network.

The second motivation is to reduce the number of transmissions without compromising on the quality of sensed data. This has been achieved using compressive sensing. Compressive sensing in wireless sensor nodes has overcome the rigid sampling procedure to acquire sensed data during data acquisition process. The sensed data obtained for transfer at the source can be represented in a sparse manner and recovered at the destination. The data forwarding from source node to the neighbor nodes is based on correlation. In short this work would focus on compressive sensing and data forwarding to reduce the sensing cost and the number of transmissions and hence achieve energy efficiency.

The third motivation focuses on reduction in observation cost using compressive sensing. The data forwarding is decided based on the capacity of intermediate nodes and available bandwidth resources. In order to effectively transfer the compressed data there should be minimal number of sensors to ensure connectivity to the sink. This is termed as the communication constraint. Connectivity in wireless sensor network defines the possibility, an intermediate node receives transmission and forwards the same to the sink.

A bottleneck in the wireless networks occurs due to the varying rates of data traffic. It can be measured and estimated using machine learning techniques. The designed communication protocol should determine whether to forward data or to refrain from transmission based on the available network resources.

1.3 Problem Statement

The first problem concerns the estimation of sensor lifetime within the deployed sensing field (terrain). The major prospective associated with sensor node lifetime is energy. Two criterions in a wireless sensor network can cause energy holes. The first criterion is cluster head selection process and uniform cluster sizes. The second criterion is when the sink is static. In the first criterion clustering: The elected cluster head acquires data from cluster members and it curtails the direct transmission from cluster members to sink thus reducing the number of transmissions. Cluster head node due to subsequent transmission may drain its resources of energy and bandwidth faster. Hence an alternate cluster node as a cluster head has to be nominated at each round to retain the connectivity. Usage of identical size clusters in the network may cause energy drainage of those clusters near the static sink due to continuous forwarding of data. Hence addressing the issues of cluster head selection and uniform cluster sizes in the network is very important. The second criterion is usage of static sink. By using multihop –communication all the cluster head nodes forward their data to the static sink. Consequently, all the nodes near to the sink deplete their battery level resulting in hotspot problem. Usage of mobile sink along with static sink can subsequently increase the network lifetime. An efficient cluster head selection process with unequal clustering and dual sink can significantly reduce the energy consumption in the network.

The second problem focuses on acquisition of sensed data, data forwarding and recovery at the sink. In the previous work data aggregation is performed based on the Nyquist sampling theorem. The data aggregation method suggests, sampling frequency should be twice as greater as the largest frequency at source for proper recovery of the signal. This process results in redundant data to be produced. Hence the further algorithms focus on compressive sensing to reduce the redundant data or data aggregation cost. The sensor node generates a huge amount of sensor readings and the

readings have to be transferred to the sink. In compressive sensing process the network data generated can be represented in a sparse manner using the spatial and temporal coordinates. Thus the sensed sparse data with fewer measurements obtained by the process of compressive sensing can be exactly recovered at the sink without any losses. Once compressive sensing is performed the problem occurs during data forwarding when all the co-located sensors transmit simultaneously leading to collisions. The process burdens the wireless network. So the proposed algorithm should focus on data forwarding from source node to any one of the neighbor nodes based on correlation thereby reducing collisions and number of transmissions.

The third problem focus on the reduction in observational cost using compressive sensing. It also emphasizes on reduction in transmission cost by considering the link capacity of intermediate nodes and network resources during data forwarding. Each sensor node has several intermediate forwarders capable of relaying the networked data to the subscribed sink. The magnitude of data at the intermediate forwarder changes the feasibility of route and its data forwarding capability. The link quality is one of the significant metrics, ensuring the packet forwarding from the sensor to the intermediate forwarder and from the intermediate forwarder to the sink. Any packet loss will burden the recovery process causing deviation from the conventional route. Hence the proposed algorithm should analyze the link capacity of intermediate nodes and the available network resources for data forwarding.

1.4 Research Aims and Objectives

The main aim of this research is to design an energy efficient data acquisition and transmission methods in wireless sensor networks. To fulfill the aim we have drawn the following objectives for this work.

1. To design energy efficient unequal clustering algorithm with dual sink to improve the network lifetime of sensor nodes within the entire topology of the sensing field. The reduction in transmission burden on the sensors is achieved by suitable cluster head selection and the operation of dual sink. Mobile sink acquires data at appropriate intervals reducing the energy overhead incurred during transmission.
2. To develop an intelligent compressive sensing-based algorithm to reduce the number of transmissions and attain energy efficiency in sensors. The co-located sensor transfers data when carrier sensing threshold of communicating entities exceeds the desired value of spatial and temporal coordinates. Thus faulty links are denied from transmission.
3. To design a unified framework of compressive sensing with perceptron-based forwarding to attain energy efficiency. Data transmission is based on the available network resources and forwarding capability of intermediate nodes.

1.5 Contributions

The contribution of clustering algorithm is the formation of unequal size clusters, cluster head election process and usage of static and mobile sink (Dual sink) in the network. The unequal size clusters and dual sink reduces the energy holes being created. The cluster head election is based on the energy of sensors nodes thus achieving load balancing. On the whole clustering and dual sink improves the lifetime of the sensors.

The second contribution is reduction in transfer volume of sensed data without compromising its Quality of Service (QoS) and attaining energy efficiency. The focus of this algorithm is mainly on data forwarding from source node to the present forwarder nodes. The data forwarding process is associated with the correlation coefficient, spatial and temporal coordinates of source and forwarder nodes. Formation of Rademacher matrix using information of spatial and temporal coordinates and data forwarding from source node to forwarder node is based on the Rademacher matrix values. This framework preserves energy consumption providing scant transmissions.

The third contribution is to reduce the observation cost and transmission cost. This is achieved by compressive sensing and through efficient data forwarding techniques respectively. In the proposed framework data forwarding is completed bearing in mind the network status and congested links. The framework tackles two scenarios for data forwarding: when the network resources are abundant and scarce. In both the scenarios data forwarding from source to sink through intermediate nodes is executed depending on the desired data forwarding threshold. The fusion rules for forwarding in order to overcome channel impairment are processed at the intermediate nodes using perceptron-based forwarding. For concreteness, in transmissions inference is made by the source node as to whether to forward or not to forward.

1.6 Outline of the Thesis

This thesis is divided into six main sections comprising of: i) introduction; ii) literature review; iii) Clustering algorithms with dual sink; iv) Intelligent Neighbor Aided Compressive Sensing – INACS v) Compressive Sensing with Perceptron based forwarding v) Conclusions

Chapter 1 **INTRODUCTION**: This chapter provides an overview of the research. The motivations and problem statements are discussed. The problem statement leading to research aims and research objectives are detailed. Finally the contributions are clarified.

Chapter 2 **LITERATURE REVIEW:** A wide range of literature review is presented in this chapter. The thesis deals with basic aspects of wireless sensor networks, energy consumption in WSN, data gathering strategies, Data acquisition in networks, impact of transmission cost and traffic intensity in available network resources, influence on network lifetime, data prediction approaches, compressive sensing techniques, sparsity and recovery in compressive sensing, channel estimation in compressive sensing, routing protocols in compressive sensing and finally computation capability of compressive sensing are discussed.

Chapter 3 **CLUSTERING ALGORITHMS WITH DUAL SINK:** This chapter examines the two-energy efficient unequal clustering algorithms with dual sink. In line with research objective 1). The chapter examines the two protocols developed where efficient clustering of sensors and sink operations are performed to acquire data. ECH-DUAL focus on inter-cluster and intra-cluster routing using dual sink. EAUC-DUAL is the further extended algorithm describing a new cluster head selection method with dual sink. Previous works on clustering algorithms, mobile sinks and static sinks are also presented. Simulation of the two protocols is performed using MATLAB. The number

of alive nodes, number of dead nodes and average residual energy are evaluated. Further the results of simulation are validated using Cox regression analysis.

Chapter 4 **INACS - INTELLIGENT NEIGHBOUR AIDED COMPRESSIVE SENSING:** This chapter explains compressive sensing and data forwarding method from source to sink in line with objective 2). The proposed INACS algorithm addresses compressive sensing, data forwarding and recovery at the sink. The algorithm elaborates the process of data forwarding based on correlation to achieve reduction in number of transmissions thereby preserving the energy of nodes. Previous works on compressive sensing techniques are discussed. Simulations are carried out to analyze number of transmissions, energy consumption and reconstruction error of the algorithm. The results are validated using time series analyses (ARIMA model) and Generalized Linear Model (GLM).

Chapter 5 **COMPRESSIVE SENSING WITH PERCEPTRON BASED FORWARDING:** This chapter explores compressive sensing, data transfer and recovery at sink satisfying objective 3) The proposed perceptron-based framework decides on data forwarding considering the link capacity of intermediate nodes and available network resources. In the framework POR data forwarding process is opted for high traffic intensity and PRMTI data forwarding is chosen for low traffic intensity. Previous works on machine learning techniques and traffic intensity are also elaborated. Simulations are performed on residual energy, energy consumption, number of transmissions, number of hops and sensing periods. The MATLAB results obtained are validated using time series analysis and expert modeler analysis.

Chapter 6: The overall summary of this work is concluded in this section and future works are discussed.

CHAPTER 2: LITERATURE REVIEW

2.1 Fundamentals of Sensor Networks

This chapter presents previous work on data acquisition and data transmissions to sink in WSN. The sensor networks emphasis on the sensing range, the communication range, and the estimation of network resources, based on which the channel estimation, energy consumption and sensor lifetime is analyzed. The communication architecture of sensors deployed within an area known as sensing field should detect the event and transfer the information to the sink. The first step of transmission is sensing followed by communication and finally the processing of data (Akyildiz et al., 2002). Sensor observations with spatial and temporal patterns have been obtained in two ways: via a single location with a single sensor or by monitoring multiple locations with several sensors, being deployed to monitor the sensing field. In the case of multiple sensors depending on the application used redundant observation can be curtailed from transmissions (Vuran et al., 2004).

Sensing rate is discussed to achieve reliability in monitoring an area within the sensing field. The sensing rate is determined by the "signal dimension" divided by the "number of sensors within the sensing field". A signal dimension has been estimated for continuous and discrete models in order to determine the sensing capacity with minimal distortion (Aeron et al., 2006). Network data processing architecture can be divided into two types depending on the target phenomena. The first is distributed sensing followed by in-network data processing (Chen et al., 2006). The second is distributed sensing from sensors processed at a fusion center or sink, known as centralized processing (Cetin et al., 2006). Sensors obtain information about particular phenomena and report it to the observer based on the desired application. The transfer path between sensor and observer is through a network protocol.

2.1.1 Data Gathering and Transmission Strategies

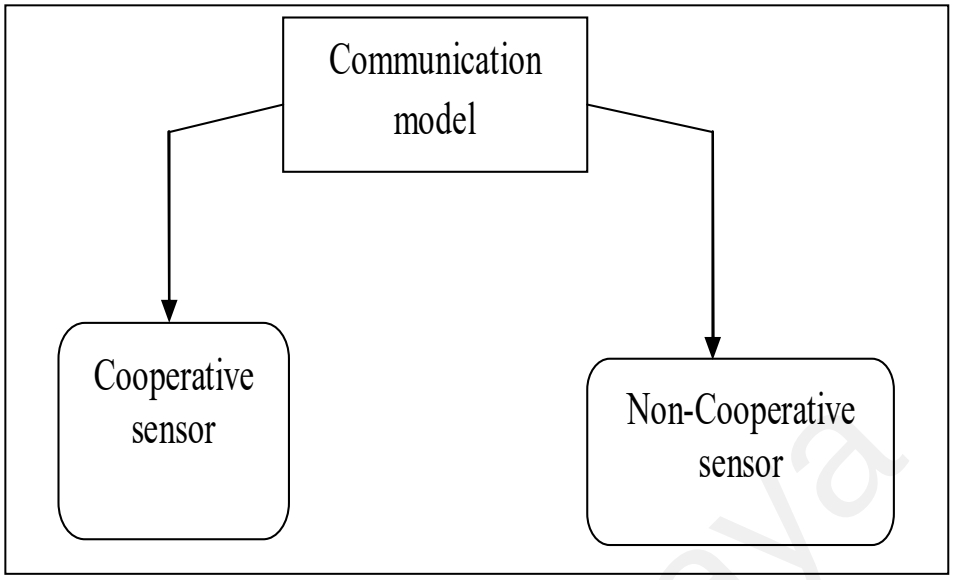

Figure 2.1: Communication model with networking of sensed data

Two communication models to satisfy the network protocols are the "cooperative model" and the "non-cooperative model" (Tilak et al., 2002). The cooperative model facilitates the sharing of information (sensed data) among sensor nodes. The non-cooperative model restricts the sharing of information. Figure 2.1 shows the communication model involving the transfer of data, as discussed in (Tilak et al., 2002). In our work co-operative sensors are used in all the proposed protocols.

Model-driven data acquisition explains the non-uniform cost incurred in the transfer of sensed data to the sink. For nodes nearer to the sink the acquisition cost is less when compared to nodes farther away from sink. The discussion reveals, rather than increasing the amount of expensive radio transmissions, appropriate storage and querying should be used to curtail the communication cost (Deshpande et al., 2004). The two most widely used models for data aggregation are "push" and "pull" (Romer & Renner, 2008). In the scenario of "push" process, reporting is made to the sink only when there is a deviation from the expected phenomena. "Push" disseminates information analogous to a broadcast. Pull-based operations acquire data from sensors using anchor nodes and transmit the data to the sink.

The performance of both the data aggregation mechanisms varies with topology and the associated traffic patterns (Romer & Renner 2008). The taxonomy of gathering data in a wireless sensor network for data fusion has been classified as "routing-driven", "fusion-driven" and "coding-driven" (Luo et al., 2009). Routing-driven approaches involve the aggregation of data for delivery to the sink with an emphasis of various performance metrics. Coding-driven approaches encompass the compression of data (source coding). Fusion-driven approaches perform redundancy removal in sensed data being denoted by an average value (Luo et al., 2009). The fusion rule for sensor networks with multi-hop strategies has been discussed using complete knowledge and incomplete knowledge of the channel status (Lin al., 2005). The likelihood ratio-based fusion rule is feasible for a low signal-to-noise ratio (SNR) environment. The information transmission is performed using the decision rule and channel statistics to achieve optimal performance (Lin et al., 2005).

2.1.2 Energy Consumption

The distribution of energy is a significant metric and can be analyzed by using the first order radio model (Heinzelman et al., 2000). Free space or multi-path fading channel models are utilized based on the distance between transmitter and receiver. The cluster member and primary cluster head (PCH) nodes consume E_{Tx} amount of energy to transmit l bits of packet over distance d with E_{Rx} amount of energy for reception where E_{elec} is the dissipated energy in a circuit and $\varepsilon_{fs}\ and\ \varepsilon_{mp}$ are the free space and multi-path fading parameters respectively. The d_0 denotes the distance at the intermediate node.

The energy dissipation model is defined in equation 2.1 and 2.2.

$$E_{TX}(l,d) = \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2, & d < d_0 \\ lE_{elec} + l\varepsilon_{mp}d^4, & d \geq d_0 \end{cases} \qquad (2.1)$$

$$E_{RX}(l) = lE_{elec} \qquad (2.2)$$

Systematic power analysis is an important factor in determining the bottlenecks in sensor nodes (Raghunathan et al., 2002). In addition to sensing and communication process, sensors also perform the operation of forwarding data from other nodes to reach the sink. Any energy drainage will lead to discontinuity in transfer of information. Wireless sensor nodes with limited energy and communication range usually exchange information with nearby nodes. Therefore by decreasing the volume of transmissions, minimal depletion of energy has been achieved (Mahfoudh & Minet, 2008). The long-time operation of wireless sensors requires proper execution of an energy budget. The energy operation can be classified based on the sensing subsystem or the network subsystem (Anastasi et al., 2009). The sensing subsystem signifies the number of samples it obtains for the event. The network subsystem emphasizes on the operation of nodes and the design of protocols for data gathering and transfer.

This thesis focuses upon both sensing and network subsystems. In the sensing subsystem the obtained data can be represented in a sparse manner for recovery. Network subsystem focuses on reduction in the transmission of sensors during data transfer.

2.2 Data Acquisition in Sensor Networks

In distributed sensor networks the transfer of sensed data can be managed in a centralized manner, distributed manner or a combination of both hybrid approaches (Chong & Kumar, 2003). Distributed sensing yields a higher SNR compared to centralized sensing when the signal of interest is unknown (Estrin et al., 2001). However in most scenarios the sensed data is transferred from distributed sensors for processing at the sink. Data dissemination with direct diffusion has been discussed (Intanagonwiwat et al., 2003). The first step is concerned with data dissemination where the sink propagates the desired data of interest to the sensors and is aware of the application. Secondly a gradient setup has been initiated where based on the interest neighbors perform local interaction. Finally a path reinforcement policy is realized and periodically updated for the transfer of data. In sensor data collection process, there are three states namely: deployment stage, control message dissemination state and data delivery state. The deployment stage ensures the coverage and connectivity measures of sensors. The control message dissemination state tries to reduce the transmissions and maintain the residual energy. The data delivery model defines the transfer mechanism between the sensor and the sink with no compromise on the metrics Quality of Service (QoS) (Wang & Liu, 2010). The relationship between coverage and connectivity in states' transmission range (R_t) should be greater than or equal to two times the sensing range (R_s) as in equation 2.3 (Zhang & Hou, 2005).

$$R_t \geq 2 R_s \qquad (2.3)$$

In-network data aggregation refers to the process of sensing (gathering of information) and routing in multi-hop communication, where intermediate nodes subsequently reduce the resources consumed resulting in elongation of lifetime (Fasolo et al., 2007). This can be coarsely divided with the objective of size reduction in the transmission of information or without size reduction in transfer of information.

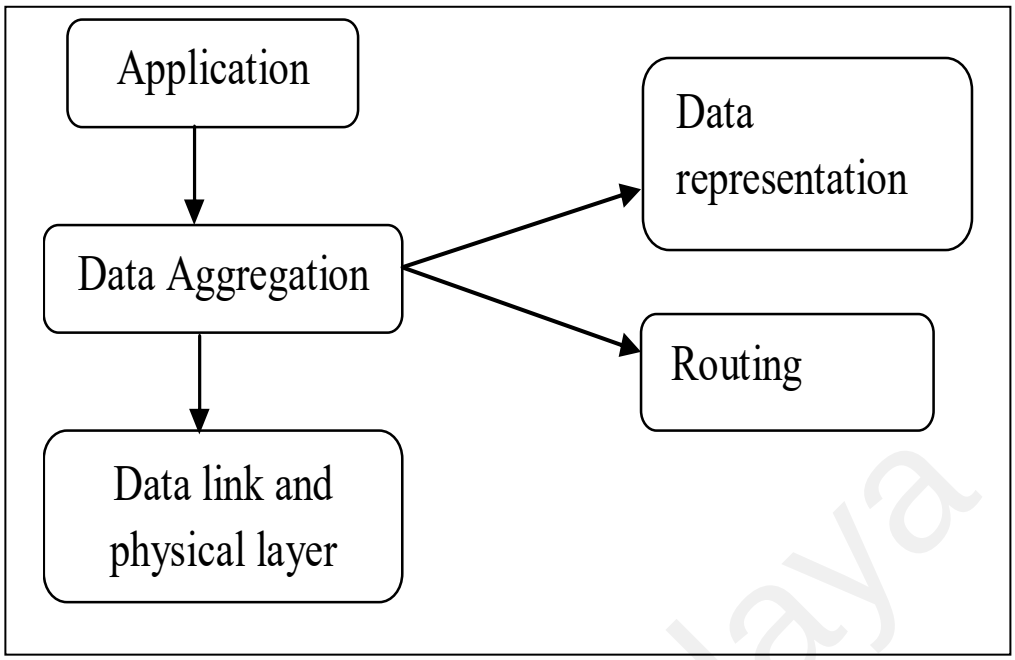

Figure 2.2: In-network data aggregation process

Figure 2.2 presents the description used for network data aggregation as in (Fasolo et al., 2007). Figure 2.3 presents the description used for data centric approach. In the data-centric approach data coming from multiple sources are aggregated at an intermediate node and transferred to the sink (Krishnamachari et al., 2002).

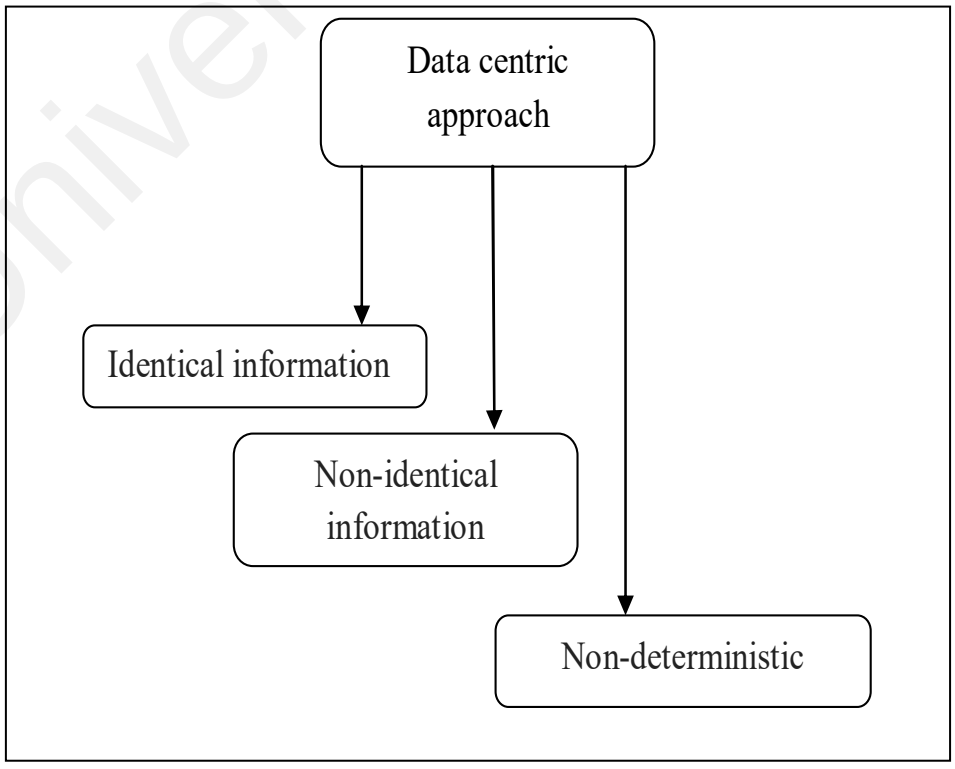

Figure 2.3: Data centric approach for information transfer

This approach relies on topology and information from sources. The sources can transfer identical information, non-identical information and non-deterministic information by imposing a burden on the intermediate forwarder in order to discriminate data. In this dissertation an emphasis on sensing the network data is performed in two ways: performance of data acquisition at sampling rate and performance of data acquisition below the sampling rate using compressive sensing technique.

The greedy aggregation mechanism suggests prior estimation followed by aggregation, based on the topology provides better path sharing and considerable energy consumption (Intanagonwiwat et al., 2002). The aggregation process has been discussed based on queries. Depending on the sensory data values Structured Query Language (SQL) queries are assigned an aggregation function (Madden et al., 2002). In the sensory network query system the adaptive sample rate and query lifetime depends on the nature of the application. In addition the query processing workload hang on the routing tree for optimization. (Gehrke & Madden, 2004).

Model-driven data acquisition explains. the locally computed sensor data can be reported depending on its predictions of the model. However when there is a deviation due to an event in the sensing field it should be reported in an appropriate interval (Raza et al., 2012). Network stack cross-layer approach uses a data link layer and a network layer to improve the network lifetime metrics of sensor networks through tuning appropriate parameters. Tuning parameters mainly are based on segregating the control operation and the data traffic (Raza et al., 2015). Literature review of sensing ranges and impact of routing protocols have been discussed in 2.2.1

2.2.1 Sensing Ranges and Network Lifetime

Adjustable Range Set Covers (ARSCs) have been proposed for duty cycle operations. ARSCs avoid redundancy while sensing a target in centralized and decentralized topologies thereby elongating network lifetime. The sensing range

adjustment should be coupled with the communication range for attaining energy efficiency. The transmission range depends on the frequency of data transmission in a specified time period of rounds (Cardei et al., 2005). However the discussion is suitable for discrete targets within the sensing field (Lu et.al 2009). The distributed lifetime coverage optimization (DiLCO) protocol where a subset of sensors alone is activated in the sensing phase has been proposed. The leader node located within the sensing field determines the over-coverage and the under-coverage areas (Idrees et al., 2006). As the problem with selecting a leader within the sub-region is difficult the duty cycle operation needs to be incorporated for monitoring the activity status of nodes.

The sensor coverage depends on the type of application (critical or non-critical) and the connectivity depends on reporting interval. Specific attention has been given to the use of a queuing model ensuring connectivity between source and destination (He et al., 2011). Full-area coverage optimization has been discussed for extending the lifetime with energy metrics (Yang et al., 2014). The minimum weight value for covering the full area is transformed into a discrete coverage ensuring the sensing threshold based on the sensor location. The attribute used for the sensing threshold is based on the approximation ratio (the rate of data aggregation divided by the rate of optimality)

2.2.2 Data Aggregation with Queries

In the clustered aggregation technique (CAG) within a sensing subfield only one value is transmitted by the cluster to decrease the number of transmissions. The error threshold is provided by the cluster for calculating the accuracy (Yoon & Shahabi, 2005). Improvements to the cluster aggregation scheme have been proposed where the spatial and temporal coordinates are considered in (Yoon & Shahabi, 2007). Two modes have been discussed according to environmental conditions namely “interactive” and “streaming”. In the interactive mode, one set of responses has been provided for a single query suitable for dynamic topologies and varying data rates. The streaming mode is

provided by the cluster head for a query suitable for static networks with no abnormal readings (Yoon & Shahabi, 2007). The curve query-based approach has been developed to monitor the continuous change occurring in the environment. It has been accomplished with a time-continuous curve associated with the retrieval of data. Depending on user needs the query is initiated and data may be procured for processing (Cheng et al., 2014).

2.2.3 Data Gathering via Sensors with Spatial Correlation

This subsection focuses on the extent of network data correlation in spatially correlated sensors. Distributed source coding (DSC) (Cristescu et al., 2005) exploits spatial correlation denoting the transmission structure to be optimized. The underlying transmission structure depends on the production rate of information and path weight. The coding module used in DSC is based on Slepian-Wolf coding. Distributed transform coding with wavelets is discussed in sensor networks (Wagner et al., 2006). The tree-based wavelet suggests, "breakeven points" are important for the transformation of communication cost. These breakeven points depend on network configuration parameters such as the number of sensors and transmission ranges determining the multi-scale transformation.

Bit hop metrics is proposed with regard to spatially correlated sensors and graph theory to calculate the number of bits transmitted per cycle (Pattem et al., 2008). The discussion highlights the joint need for compression and routing. The in-network modeling of sensory data is discussed with kernel-based regression. The behavior of spatial-temporal processing through correlated data can be predicted and communication cost can be reduced (Guestrin et al., 2004).

2.3 Discussions on Network Resources

This section deals with the usage of network resources from transmission cost to traffic intensity in line with energy consumption and network lifetime of sensor nodes. The impact of sensor nodes on a single radio channel requiring bandwidth usage is critical. Thus reducing the redundancy of sensor reading in a distributed manner is vital (Pradhan et al., 2002). Spatial techniques focus on working with a group of sensors within the sensing field while temporal techniques concentrate on individual sensors and their time domain.

2.3.1 Impact of Transmission Cost

Distributed computing in wireless sensor network communication occurs at the wireless interface by determining the choice between forwarding or not based on available bandwidth. Hence the communication protocol developed should meet the requirements of communication bandwidth meant to be scarce, long and incoherent with an enormous data rate produced by the sensors (Yick et al., 2008). The wireless channel exhibits power law decaying factors such as reflection, scattering and diffraction leading to attenuation. This causes the channel impulse response to be sparse.

The recovery of such signals in a channel is especially difficult (Choi et al., 2017). The pseudo probabilistic model has been discussed in terms of reducing the communication cost (Singh et al., 2018). It explains, within a sensing field the joint sensing probability is much higher than the predefined threshold. The detection probability decreases as the SNR increases. However the approach is difficult due to the topology of sensor nodes and the formation of the Johnson circle with a minimal failure point.

2.3.2 Impact of Traffic Intensity

Fairness in transmissions is computed in wireless sensor networks for one to many communications. In this case fairness denotes an equal number of packets to be transmitted towards the destination (Ee & Bajcsy, 2004). This is achieved with probabilistic selection or epoch-based proportional selection depending on the calculations of the available bandwidth and the size of the sub-tree (Ee & Bajcsy, 2004). Probabilistic selection based on equal probability has been allocated to all nodes. Epoch-based proportional selection relies on time epochs for transmitting an equal number of packets.

The QoS in wireless sensor network based on the application requirement has been classified into two types (Chen & Varshney, 2004). It may be end to end (observed data processed at the sink) or not end to end (intermediate nodes perform the processing) or it may even be tolerant to delay or intolerant to delay. The optimality that the flow of information to sinks in wireless sensor networks can be either based on reliability or timeliness is stated in (Felemban et al., 2006). The discussion explains, local decision criteria are needed at intermediate nodes in order to obtain the desired QoS. Two models for characterizing traffic in wireless sensor nodes with latency are discussed namely: regular data reports (RDRs) and urgent event notifications (UENs) (Wang & Wei, 2016). In the first scenario of RDR sensors transmission from source to sink occurs periodically at a constant rate where latency is permissible. In the second scenario of UEN, sensors transmission from source is unusual and unexpected and the reporting interval should match in terms of the appropriate latency.

Regular reporting of link losses between wireless sensor nodes and the sink can consume more overheads. Hence segregating certain links with high loss and providing end-to-end transmission are important for sensors pertaining to local conditions (Nguyen & Thiran, 2006). Time-varying networks with Oseledec's theorem denoting

the convergence speed of the algorithm is discussed (Denantes et al., 2008). The discussion reveals the contraction rate depends on the metrics of energy consumption and the number of transmitted messages (Denantes et al., 2008). In priority-based link scheduling the major constituents of priority in scheduling are traffic load, the number of nodes attached to the parent node and finally the hop count with regard to the sink (Huang & Soong, 2017).

2.3.3 Sensor Centric Approaches

The sensor-centric approach is proposed where nodes decide to "send" or "not to send" based on the informative observation obtained within the sensing interval (Rago et al., 1996). Similar approach has been discussed considering the reduction in communication cost. The work focuses on peripheral sensor nodes in small-scale fading channels within the sensing field (Jiang & Chen, 2005). Transmission in wireless sensors discussed along with mutual information rates and censoring scheme can reduce the observation cost and transmission cost by appropriately scheduling the censoring approach and the probability of sleeping (Yamasaki & Ohtsuki, 2005). A reduction in data volume within the sensing field using a subset of sensors is discussed where certain sensors provide transmissions to the fusion center based on two criteria: the maximum likelihood estimate (MLE) for deterministic networks and the maximum posterior probability for random networks (Msechu & Giannakis, 2011).

2.3.4 Influence of Energy Consumption

The radio frequency (RF) component used in sensor nodes for information flow defines, the interference rejection is lesser for short-range transmission with a lower duty cycle and a lower data rate (Bult et al., 1996). It is necessary to analyze the power consumption using a step-by-step approach after which optimization can be performed (Raghunathan et al., 2002). The two pivotal components involved in power consumption are RF components and electronics components. RF components

characterize the metrics of transmission distance and the type of modulation used. Electronics components deal with the consumed power during circuit operations (Raghunathan et al., 2002). Communication energy is discussed considering the data link layer and the physical layer (Shih et al., 2001). The computation of communication energy explains, a consolidated sum of energy is needed to transmit data by radio in addition to the energy needed to encode and decode data (Shih et al., 2001). In general network-wide energy has multiple dimensions where individual sensor contributions are high.

Energy efficiency in sensor networks is a consolidated metric involving all of the open system interconnect (OSI) layers (Jones et al., 2001). It also ensures, there should be no bias towards communication or computation components when seeking to achieve the desired energy (Jones et al., 2001). Two levels of efficiency in wireless sensor nodes have been discussed namely the best-effort model and the reliability model. In the best-effort model there is no guarantee of data delivery whereas the reliability model guarantees it. The energy per bit discussion was made in relation to these two models with an emphasis on the data link and network layers (Cao et al., 2006).

Systematic power consumption has been realized in small-scale fading (Khojastepour & Aazhang, 2004). Fading channel capacity explains the average power and peak power to be calculated in order to achieve a certain data rate. This depends on fading coefficients with gradually increasing transmissions from zero power to constant power (Khojastepour & Aazhang, 2004). The impact of energy consumption is higher in fading channels compared to non-fading channels (Fazel et al., 2013). Maximum equivalent power consumed by a power amplifier for a single and multi-hop communication has been discussed where the radio environment, the drainage characteristic of the power amplifier is to be considered for minimal power usage

(Wang et al., 2006). Bidirectional Medium Access Control (B-MAC) comes with a bidirectional interface configured according to the workload condition (Polastre et al., 2004). BMAC uses multi-hop routing typified by the signal strength from middleware (Tiny OS). Energy consumption prediction has been performed with network-wide synchronization of all nodes (Polastre et al., 2004). Energy preserving has been discussed in association with communication and computation (Kumar & Lu, 2010). The energy saved in offloading can be denoted by below equation 2.4 as in (Kumar & Lu, 2010).

$$E_{offloading} = \alpha C_i - \frac{\beta}{B_w} D \qquad (2.4)$$

The terminologies C_i denote the communication data and D denotes the amount of instructions in transfer of data to the sink. B_w denotes the bandwidth. Finally α and β represents power consumption which is a constant value of the node and sink or cloud. Five different taxonomies have been proposed for energy efficiency (Rault et al., 2014). The first is radio optimization deals with assigning suitable transmission power, antenna design and suitable modulation techniques. The second is data reduction either in the transmission or in the acquisition of signals. The third is the sleep-wake scheduling of nodes and MAC protocols. The fourth is energy-efficient routing via clustering and mobile sinks. The fifth involves battery repletion using energy-harvesting schemes.

2.3.5 Energy Harvesting

Harvesting architecture is an energy transfer process between the environment and the sensor nodes. The modelling approach can be at high level or at low level. The high level indicates the system module as a whole circuit with all its modules. The low level indicates the specific modules used for recharging an electronic circuit (Bader et al., 2014). The general taxonomy of energy harvesting system with wireless sensor node has been discussed. The three main components are harvesting source, harvesting system and load. The major sources of harvesting are solar, thermal, wind, radio

frequency and Electromagnetic waves. Harvesting system provides storage and power management capability. Harvesting Load denotes the sensors and its transceiver used for data transfer (Sherazi et al., 2018). Energy harvesting schemes rely on surrounding environment. The diurnal changes of solar radiation cause alterations in transmission radius of sensor networks. Energy Conserving Transmission Radius Adaptive (ECTRA) has been proposed where nodes can alter their transmission radius depending on the environment. Thus hot spot due to energy imbalance are reduced near the sink and reduced network delay has been achieved (Ju et al., 2018).

2.3.6 Influence of Network Lifetime

A node is said to be alive when it performs two operations. The first is sensing by sensors and reporting to the sink while the second is simply relaying information where the node is active but performs neither sensing nor relaying (Bhardwaj et al., 2001). On the contrary a node is defined to be dead or inactive when it loses the coverage. Source nodes should characterize the feature of minimal energy in forwarding to the relay or the sink. Lifetime can be estimated by two means. The first is when sensor nodes run out of energy. The second is when a fraction of sensors run out of energy or while a network partition occurs during forwarding. The most significant factor lies in the underlying network model being deployed (Chen & Zhao, 2005). According to discussions in (Shah & Rabaey, 2002) the low energy path may not be ideal for prolonging network lifetime. Hence alternate paths from a node must be designed without the saturation of resources (Shah & Rabaey, 2002). Network lifetime states, the metrics for maximizing the lifetime should also match the maximization of the information flow. The intermediate node has a vertex capacity for forwarding to its neighbor in the vicinity or directly to the sink. However the simulation in this work considers a slice model with equal distance (Jarry et al., 2009). The issue of network lifetime with varying transmission ranges depending on their location has been

discussed (Perillo et al., 2004). In this scenario the farthest nodes in the sensing field communicate via multi-hops imposing a burden on midfield and nearer nodes. Thus it causes an imbalance in network-wide energy leading to drainage of the energy from nearer nodes (Perillo et al., 2004). The network lifetime and deployment strategies is discussed in (Cheng et al., 2008). The main focus of the deployment scenario concerns traffic type (uniform or non-uniform), sensors (homogenous or heterogeneous), energy (uniform or non-uniform) and increasing the number of sinks (static or mobile). Hence the corroboration of extended lifetime is contingent with varying metrics and applications.

The ageing problem of sensor nodes in sparse and dense networks has been discussed. In a sparse network an increase in hop level mainly leads to a reduction in network lifetime. In a dense network depletion of energy occurs due to high workload involved in forwarding at the first hop node (Lee et al., 2008). The impact of traffic load associated with the radio model and its proximity to the sink has been discussed using per node traffic. The discussion on the energy budget states, the maximum amount of energy incurred during transmission reduces the operational lifetime of sensor nodes (Chen et al., 2009). Extending lifetime is discussed by assigning a weight value to an individual node and calculating the residual energy (He et al., 2011). Accordingly a subset of sensors with minimal energy activation is selected in order to prolong network lifetime. The work discusses on reducing data by correlating the spatial and temporal properties of sensor nodes in a sensing field, thus preventing the drainage of energy (He et al., 2011).

2.3.7 Data Prediction Approaches

Wireless sensor network performance with data mining approaches has been discussed (Mahmood et al., 2013). The traditional data mining cannot be incorporated in wireless sensor nodes because the data type is dynamic (changing fast in nature), the

data flow is continuous and the response time does not complement the real-time application. There are several data prediction approaches depending on the application and the time interval deciding on the next data points (Dias et al., 2016). The major taxonomy of predictions has been classified into single predictions and dual predictions. Both schemes focus on generating the prediction model either at the cluster head or at the sensor nodes.

2.4 Compressive Sensing

In compressive sensing objects (signals) are represented by vectors and can be reconstructed with fewer measurements using near-optimal algorithms (Donoho, 2006). The process of transmission for a sampled signal with fewer measurements and recovery needs can be obtained with compressive sensing (Candes et al., 2007). This is because two of the proposed works are concerned with compressive sensing involving the acquisition of sensed data and it varies from the conventional data acquisition. The preliminary stage of compressive sensing is discussed in the following sections.

2.4.1 Preliminaries in Compressive Sensing

The characteristics needed for compressive sensing at the source node are sparse bases, the representation of measurements and a sensing matrix. From the receiver recovery viewpoint we should follow the Restricted Isometry Property (RIP) and minimal reconstruction errors. This communication methodology at the source and the receiver should match the dynamic environmental conditions prevailing between the source and the sink. The main principles behind compressive sensing are sparsity and incoherence (Candès & Wakin, 2008). The term *sparsity* in the case of a sensor has been represented in two ways: the first concerns the sensor's topological coordinates while the second concerns the estimated sparsity of sensed data in sensing field (Marvasti et al., 2012).

In this research exploiting sparsity is executed by estimating it within the sensing field rather than the location coordinates of the topology. The signal must be sparse enough to enable compressive sensing. Sparsity of a vector states, the number of non-zero entities is smaller than the original dimension of the vector. Generally the l_0 norm or l_1 is used to check sparsity of the signal (Candes et al., 2006). The sparsity has been discussed with the equation 2.5 with l_1 norms.

$$m \geq 4s \qquad (2.5)$$

It states the number of measurements (m) should be greater than or equal to four times the non-zero coefficients (s) (Candes et al., 2008). Maximal incoherence should be present between the sensing matrix and the sparsity basis (Candès & Wakin, 2008). The property of coherence in compressive sensing suggests, the two columns' correlation is high. If this is the case it is difficult to analyze the energy from where the signal has come (Candes et al., 2011). Hence to reduce the burden of post processing it is necessary to address the property of incoherence. The relationship between sparsity and incoherence indicates the process of reconstruction for a sparse signal cannot be achieved with a fixed measurement matrix (Candes & Romberg, 2007). The discussion states, sparsity is a varying factor and it could be natural; otherwise transformation technique needs to be applied to achieve sparsity (Candes & Romberg, 2007).

2.4.2 Sparsity and Measurement Matrix

Compressive sensing with single and multiple measurement vectors is addressed in (Tropp et al., 2006). Single measurement vector denotes the combination of basic signals where the sparsity can be linear in nature. In multiple measurements the sparsity varies and is not found to be linear in nature. In some scenarios the signals themselves are sparse and compressive sensing can be incorporated easily (Duarte et al., 2005). Two such models have been discussed: Joint Sparsity Model-1 (JSM-1) and Joint Sparsity Model-2 (JSM-2). In the case of JSM-1 there are two types of signals those with a common sparse part and those with no identical parts. Signals change smoothly in spatial and temporal domains with high correlation exhibit the property of JSM-1. In the case of JSM-2 a vector basis is used to provide the same sparse support while the non-zero coefficients are different (Duarte et al., 2005).

Model-based compressive sensing has the advantage of structured sparsity in recovery (Baraniuk et al., 2008). Firstly it implies, the number of measurements can be considerably reduced if structured sparsity is incorporated. Secondly structured sparsity

discriminates true signal from artifacts. The sparsity of signals has a time-varying feature which changes as the signal travels from the source sensor to another sensor. The transaction probabilities of signals from slow to fast variations has been discussed along with multiple measurement vectors (MMVs) and least squares methods (Bawane & Kannu, 2019). The results reveal a change in the measurement vector together with its corresponding transaction probability achieving a higher gain.

According to the measurement matrix discussion the transformation of the sparse representation in order to recover a signal occurs through the encoding matrix. This measurement matrix can be deterministic or random. The deterministic matrix satisfies the incoherence criterion for all the signals. The random matrix satisfies the incoherence criterion for certain signals (Arjoune et al., 2018). The advantages of the deterministic matrix are desirable structure, minimal storage and can be used for fast implementation (Nguyen & Shin, 2013). However it requires prior information about the signal. Performance analysis of wavelet-based sensing matrix in one dimension has been discussed in (Yang et al., 2010). Compression ratio and sparsity level can be obtained by equation 2.6 and equation 2.7. The “k" in equation 2.7 denotes number of non-zero entity.

$$\text{Compression ratio} = \frac{\mathrm{m}}{\mathrm{N}} \tag{2.6}$$

The “N_l” denotes length of signal, “m” denotes the measurement acquired. The N denotes the dimension of the signal to be compressed at source.

$$\text{Sparsity level} = \frac{\mathrm{k}}{N_l} \times 100 \tag{2.7}$$

2.4.3 Reconstruction Accuracy in Compressive Sensing

Incorporating a suitable measurement matrix can improve the information’s theoretical limits during recovery (Wang et al., 2010). The three pivotal terms involved in determining the lower bound of sparsity for recovery are as follows: “dimension of the signal with its represented function”, “sparsity of signals” and “sparsity of

measurement" (Wang et al., 2010). The wireless channel exhibits power law decaying factors such as reflection, scattering and diffraction leading to attenuation. This causes the channel impulse response to be sparse. The recovery of signals in such a channel scenario is very difficult to be achieved (Choi et al., 2017). Thus the challenge lies in capturing the network resources and designing a proper matrix in order to ensure data recovery at the receiver.

The sparse distribution of the unknown signal should be decentralized in an adaptive and distributive manner for recovery (Di Lorenzo & Sayed, 2012). The vector considers the sparsity pattern changes overtime and has to be tracked. Hence sparsity regularization was discussed in relation to the l_0 norm indicating the number of non-zero entities to be restricted with an upper bound and known prior for recovery. Two versions of sparse signal recovery have been proposed for saturation errors namely "saturation rejection" and "saturation consistency" (Laska et al., 2011). Saturation rejection denies saturation errors and performs compressive sensing recovery. Saturation consistency implies, the recovery can be attained using convex inequalities.

2.5 Compressive Sensing in Wireless Sensor Networks

Compressive sensing can be incorporated for a single wireless sensor node or for a group of nodes (Sarvotham et al., 2005). The process usually involves a group of nodes is known as distributed compressive sensing (DCS). DCS is implemented with a multi-signal in order to reveal the joint sparsity characteristics of individual signals. A multitude of wireless communications such as channel fading and sensor node properties can lead to adverse effects. To counter these effects a proper design using the available resources are necessary. Compressive sensing is an approach for a networking community where storage, transmission and retrieval take place without prior knowledge of the signal (Haupt et al., 2008). The two features of compressive sensing are "universal sampling" and "encoding in a decentralized way".

The sampling rate in compressive sensing depends on the content of information of a signal (Mishali & Eldar, 2012). In traditional sampling the sampling rate relies on the signal bandwidth. Hence there is a considerable reduction in the acquisition of signals. The sampling process explains the union of subsets can be used in order to minimize the resources (Mishali & Eldar, 2012). The three major components used for achieving this are the analog-to-digital converter (ADC), the digital signal processing (DSP) toolbox and the digital-to-analog converter (DAC). The ADC converts the analog signal to a union sequence, the DSP toolbox performs the required signal processing tasks, and the DAC is used for the reconstruction of the signal from the received samples.

Sparse sensing explains the signal required for processing does not need to be sparse for recovery; however the compressive sensing signal needs to be sparse for recovery (Chepuri & Leus, 2016). Sparse sensing and compressive sensing both focus on data reduction. The process involved in sparse sensing is as follows: "estimation", "filtering" and "detection" which incur complexity in processing due to perturbation of

the signal in the sampling space (Chepuri & Leus, 2016). Compressive sensing data are quantized uniformly and transferred to the sink. A resourceful sink achieves high reconstruction accuracy based on processing capability, rate distortion and appropriate quantization intervals. The network infrastructure and estimating the transfer capabilities of individual sensors through an awareness of the wireless environment are mandatory (Felipeda Rocha Henriques et al., 2019).

2.5.1 Channel Estimation in Compressive Sensing

The estimation of channel features is insufficient to form the channel matrix required for the process of compressive sensing. The basic pursuit method exploits the features of time-invariant channels so as to attain minimal error in compressive sensing (Berger et al., 2010). The recovery of noisy signals has been discussed in relation to compressive sensing using two sensing channels namely: independent and identically Gaussian observations, correlated observations (Aeron et al., 2008). Sparsity has been defined as the smallest number of observations with minimal distortion and the same signal can be recovered within the sensing channel.

2.5.2 Data Aggregation in Compressive Sensing

The minimal approximation error has been obtained in compressive sensing by selectively querying certain sensors. The selection of the queried sensors depends on the coefficients of the data collected. Thus the communication cost is reduced in per packet transmission (Wang et al., 2007). The scalability of sensors with common sparsity has been discussed (Duarte et al., 2009). The results show, as the measurements of network data increase per sensor, recovery accuracy can still be attained without increasing the compressive sensing measurements. Sensor perception for recovery describes the “event of interest for recovery” and the “sampling function” are the two criteria used to measure the physical event (Yang et al., 2010).

Compressive sensing has been proposed with aggregation of data and arbitrary topologies using compressive data aggregation (CDA). In CDA sparsity is achieved with diffusion wavelets. Incorporating network partition results in spatial and matching temporal coordinates where the former achieves high fidelity in reconstruction (Xiang et al., 2013). The work discusses compressive sensing of sensed data in relation to the autoregressive (AR) model and through it the sparsity is determined (Wang et al., 2012). Based on the reconstructed data and their recovery errors, the sparsity parameters are tuned. The data gathering process discriminates between "critical" or "non-critical" data and accordingly the transmissions to the sink take place (Rao et al., 2019). In the case of critical data the prediction-based model is incorporated and for non-critical data compressive sensing is used.

2.6 Energy Consumption in Compressive Sensing

Compressive sensing varies from the general signal processing architecture used for data acquisition. Hence the energy consumption model for procuring sensed data also varies as has been discussed with regard to data acquisition and compressive sensing (DACS) (Karakus et al., 2013).

$$E_{SP-DACS} = N_o\varepsilon_{mr} + M_cN_o\varepsilon_{add} + M_c\varepsilon_{mw} \qquad (2.8)$$

Energy required for Sensor processing $= E_{sp}$
Energy required for DACS $= E_{DACS}$
Energy dissipation for memory reading $= \varepsilon_{mr}$
Energy dissipation for memory writing $= \varepsilon_{mw}$
Energy required for addition $= \varepsilon_{add}$
N_o denotes original signal acquired and M_c denotes compressed signal.

A potential physical layer for censoring and compressive sensing has been proposed (Wu et al., 2018). The sensor node decides whether to send or to refrain from transmissions using partially available knowledge of the signal. A homogenous environment has been considered for simulations and accuracy is obtained in low and medium environments.

2.6.1 Routing Protocols in Compressive Sensing

Compressive in-network data processing (CIDPS) achieves balanced network traffic in wireless sensor nodes. A load balancing feature has been achieved at intermediate nodes, considering the link status and the compression ratio. A threshold value is assigned based on the traffic scenarios (Singh & Kumar, 2018).

On-demand explosion-based compressive sensing (ODECS) has been proposed where the procedure works within the sensing field aimed at adaptation to the occurrence of events (Singh et al., 2019). ODECS mainly exploits the change in data rate for events incorporating compressive sensing and data gathering. In ODECS a protracted network lifetime is achieved rather than in hybrid compressive sensing. It

also eliminates the bottleneck load near to the sink. Compressive Sensing-based Path Reconstruction (CSPR) collates the length of the routing path and the network size in order to find feasible routes. CSPR exploits sparse paths using compressive sensing while the reconstruction is performed at each path with a small packet. Further a suitable path with minimal loss is chosen (Liu et al., 2016).

Compressive Sensing based on Packet Loss Matching (CS-PLM) discusses the packet loss reducing the postprocessing accuracy of the gathered data. Since the packet transmitted might belong to a transmission super-imposed by several nodes a multi-path backup scheme is proposed. It combines traditional forwarding and compressive sensing (Sun et al., 2018). Clustered compressive sensing (CCS) incorporates clustering and compressive sensing (Sun et al., 2018). It uses two methods for direct communication between the cluster head and the sink (D-CS) and multi-hop communication between the cluster head and the sink (I-CS). The block diagonal measurement (BDM) matrix determines the cluster size within the sensing field in CCS achieving a reduction in power consumption (Nguyen et al., 2016).

In energy consumption and optimized compressive sensing (ECO-CS) an appropriate measurement matrix with a suitable size of measurements can reduce the unwanted sampling rate and hence achieve energy efficiency. The collector node which transfers data to the sink denies transmission to the sink, if reconstruction accuracy is not guaranteed by a stopping rule (Yang & Wang, 2018).

2.6.2 Computational Capability of Compressive Sensing

Hierarchical data aggregation and compressive sensing (HDACS) has been proposed and assigning varying compression thresholds depending on the size of the cluster and data aggregation mechanisms are discussed (Xu et al., 2015). Considerable energy consumption is achieved with radio and processor models using cluster specified parameters.

Forwarding Tree Construction and Link Scheduling (FTCS) states, apart from data gathering process the sensor node decides whether or not to schedule its transmission in a decentralized way. The decentralized method relies on the physical interference model associated with its local neighbor (Ebrahimi & Assi, 2015). Topology-aware data aggregation (TADA) in static scenarios uses information topology to reconstruct the data with accuracy. A weight coding procedure converts raw sensor readings into corresponding orthogonal vectors using topological information involving minimal energy consumption. Minimal energy consumption is achieved with short weight vectors (Wang et al., a, 2019). Vandermonde matrix-based scalable data aggregation (VSDA) discusses, when the new nodes are added the measurement matrix can be reconfigured depending on topological changes. VSDA also involves a framework in where the measurement matrix can be extended based on the prior matrix (Wang et al., b, 2019).

To maximize the computational and communication aspects of compressive sensing two algorithms were proposed namely: compressive data storage probabilistic broadcasting (CStorage PB) and compressive data storage alternate branching (CStorage AB) (Talari & Rahnavard, 2016). In the first protocol the data collector or source node queries certain nodes in the transmission range in order to acquire its measurement vector. Then the pursuit method is used for compression and transmission takes place. The limitation of CStorage PB is, it depends on network topology and is prone to failure with any small changes. Meanwhile CStorage AB determines the number of next-hop forwarding nodes (two-hop) in order to contend with the network changes. This process ensures successful delivery for varying transmission ranges.

2.6.3 Wireless Network and Embedded Processing

The important aspect of wireless sensor node is sensitive operation of reporting an event within the time interval. Hence this section deals with practical implementation including control timing performance. Real time control system with MATLAB and Simulink can be extended using embedded processors. It has been discussed in the algorithms namely “Jitterbug” and “Truetime”. Jitterbug considers packet delay and provides subsequent compensation. Truetime incorporates a feature where transmission time can be assigned randomly or based on the input data (Cervin et al., 2003). Response to external events and handling the event with respect to time has been discussed with “Harvard event driven approach” where lightweight events are handled directly without the aid of software handler. The hardware systems with intelligent operation of circuits can provide minimal power than the general-purpose controllers. (Hempstead et al., 2008).

The transition from simulation software to high level abstraction named as Middleware approaches hides the complexity in hardware (Mascolo et al., 2005). The Middleware in wireless sensor node has to support the “programming abstraction”, followed by “system level abstraction”, “runtime support” and “Quality of Service” mechanism (Wang et al., 2008). Quality of Service denotes the wireless resource management and reporting. This thesis entirely focuses on the simulation of wireless sensor networks using clustering and compressive sensing algorithms.

CHAPTER 3: CLUSTERING ALGORITHMS WITH DUAL SINK

3.1 Introduction

Clustering is a significant process improving network lifetime and energy efficiency of WSN. Several techniques have been proposed in WSN so as to improve network lifetime. Compared with flat routing protocols, cluster based (hierarchical) protocols efficiently manage the sensor nodes to afford a better available route. Cluster based network reduces the overhead of communication by involving intra-cluster and inter-cluster communication. Two conditions in a wireless sensor network can cause energy holes. The first condition is when the clusters are identical. The second condition is when the sink is static. The existing LEACH algorithm uses identical clusters and a static sink. All the clusters in the network transmit their data through the cluster nearest to the static sink. So those clusters closer to the static sink become reduced in their energy level and experience the so-called hot spot problem. The same concept is applied when the sink is static. By using multi-hop communication all the nodes forward their data to the static sink. Consequently, all the nodes near to the sink deplete their battery level resulting in hotspot problem. Due to equal cluster sizes and presence of static sink, hot spot area is created. The clusters farther away are incapable to transmit to the static sink, crossing the hot spot area though they possess significant energy. To mitigate this hot spot problem mobile sink is introduced along with static sink (dual sink) for data collection in EAUC-DUAL algorithm. In the further extended ECH-DAUL algorithm in addition to dual sink, a novel cluster head selection method is proposed for unequal clustering. Thus, the two criteria of hot spot problem namely unequal cluster size with novel cluster head selection process and sink mobility are addressed in ECH-DUAL algorithm.

3.2 Previous Work on Sink Mobility and Clustering

Clustering can be an ideal solution for reduction in energy as well as reduction in transfer of messages causing problems at the cluster head node. Clustering with localizing capabilities explains the routing table reduction ensuring network scalability. However, assigning the cluster head in a predefined manner for the transfer of data can lead to node disconnection or energy imbalance (Abbasi & Younis, 2007). In the case of a single sink with a small-scale fading channel such as Rayleigh and additive white Gaussian noise (AWGN) there is a high varying non-uniform data flow from the nearer nodes to the sink. Hence protocol development should focus on the application needed to curtail the death of sensors nearer the sink (Haenggi, 2003). The nodes those can deliver data via a single hop to the sink are subjected to fast energy drainage. Hence an appropriate transmission power for delivering data from the single hop node is necessary for communication (Haenggi, 2004). Two data transport capacity metrics for reliable transfer to a single sink have been discussed (Marco et al., 2003). The first is the total transfer capability of all nodes to a sink while the second is per node transfer capability (Marco et al., 2003). Thus a single sensor death due to energy depletion can drastically affect the transfer capability.

Energy-Balanced Data Gathering (EBDG) has been discussed in (Zhang & Shen, 2008). The sensing field is divided (corona) in such way, the location of the sink is centered in the sensing field. Energy consumption within the corona (intra) is based on the amount of data within the corona and the depleted energy in its time slot. Subsequently inter-corona communication takes place in order to balance energy consumption. The problem of EBDG is the distribution of data involving dynamic metrics depending upon the application. Routing with a mobile sink can prolong sensor lifetime more than with a static sink (Luo & Hubaux, 2005). Hence a promising solution is to interpret the appropriate mobility pattern of the sink for data gathering. Path

traversal and data transfer analysis in the network topology with a single sink within a sensing field causes the sensor nodes to run out of energy. As such the conjunction of mobile and static sinks within the same sensing field alleviates sensor node death (Luo & Hubaux, 2005). Mobility pattern prediction can be introduced within the network infrastructure in a suitable environment to decrease the communication cost (Chakrabarti et al., 2003). The challenge faced by wireless sensor nodes with mobile elements is, shortage of contact detection duration due to a limited transmission range. Thus the sensor has to reliably transfer within the duration of contact detection to the mobile element. In terms of controlled mobility this should be established with a suitable path and appropriate node speed (Di Francesco et al., 2011). The sink mobility changes the process of communication by focusing on sensor nodes within a circular sensing field. Nodes within the periphery are considered for transmission (Luo & Hubaux, 2005). A mobile sink due to its movement has to establish network-wide synchronization with nodes for transfer efficiency. A dual sink - where there is a static sink and a mobile sink achieves better efficiency and the mobile sink only controls a subset of nodes (Wu & Chen, 2007). The problem with a mobile sink is it incurs more overheads through control packets due to its broadcast nature.

The gathering of data in a real-time wireless sensor networks is classified into "sink relocation" and "data dissemination to the mobile sink" (Nayak & Stojmenovic, 2007). In the process of sink relocation shorter paths have to be taken by focusing on the features of load balancing and energy scenarios. Load balancing indicates the storage of accumulated data and the receipt of new data. The dissemination of data is a combination of localization and routing. In this process the mobile sink informs its navigation pattern and sensor route data (Nayak & Stojmenovic, 2007). The criterion of clustering in wireless sensor nodes using time coordinates has been discussed in (Abbasi & Younis, 2007).

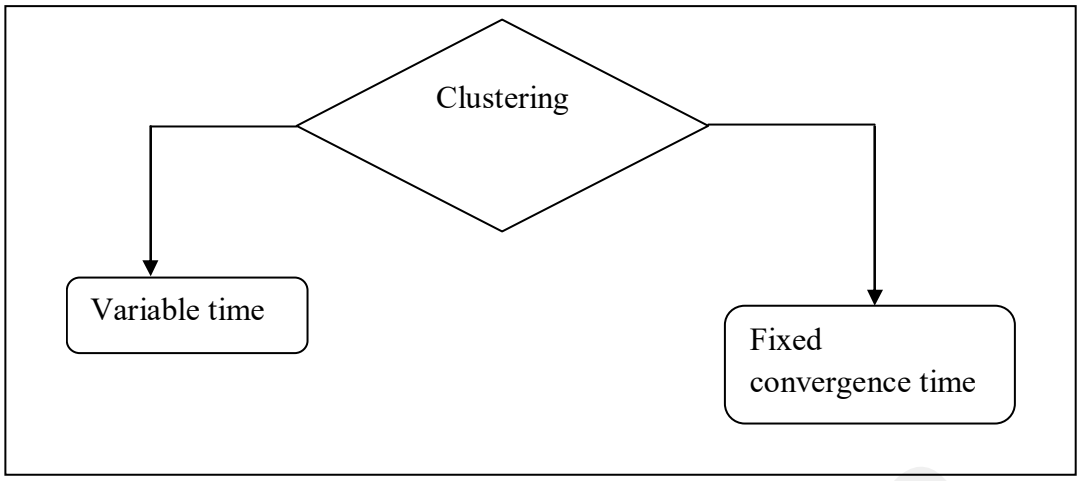

Figure 3.1: Clustering with temporal coordinates

The taxonomy of clustering is based on variable and fixed convergence time. These two forms are mainly associated with the clustering process, cluster head capabilities and clustering properties. Figure 3.1 shows time coordinates of clustering as in (Abbasi & Younis, 2007). The operational cost of transmission is high in non-clustered networks compared with clustered networks (Vlajic & Xia, 2006). In non-clustered networks transmission cost is a function of message length, the number of bits and the number of sensors deployed. In a clustered network transmission, cost is a function of the number of cluster heads. Hence cluster head selection, inter-hop distance and reduction in transmission should be achieved in clustering (Vlajic & Xia, 2006). Optimal cluster size depends on the following factors: the number of source sensors, the distance to the sink, and the spatial correlation existing between them (Pattem et al., 2008). In the Dynamic Clustering Mobile Data Collector (DCMDC) the network cluster is made optimal along with the division of the sensing field into appropriate subzones (Abuarqoub et al., 2017). The discussion on DCMDC discusses about the smaller clusters having the tendency of buffer overflow. Hence to achieve load balancing optimal cluster size is needed.

Synchronizing cluster heads with gateway communication achieves appropriate load balancing. This load balancing strategy has been obtained for varying data generation with equal and unequal loads (Kuila & Jana, 2014). However estimating the traffic load prior to cluster formation is a complex task in wireless networks. The discussion on cluster radius suggests, the remaining energy consumption and cluster head rotation are not the only metrics needed for extending network lifetime. Further it is noted, due to the impact of network lifetime on unequal clustering, clusters close to the sink should have a smaller size compared to clusters away from the sink (Liu et al., 2010). The Arbutus protocol calculates the energy consumed depending on the load. The channel state information used in Arbutus is as follows: "bottleneck link quality" is a measurement associated with link status, hop count and load. It is calculated as the ratio of relaying locally generated data within the desired timeslot (Puccinelli & Haenggi, 2008). Thus the cost incurred is calculated and premature bottleneck links are eradicated. However load disruption occurs in varying rates and cannot be easily calculated in sensors.

Minimum Transmission Energy (MTE) in forwarding might not provide an optimal solution for transmission energy (Chang & Tassiulas, 2004). The information generation rate is either fixed or arbitrary and the link capacity consisting of number of routed packets are pivotal factors in network lifetime. In the case of application-specific low-power routing, the implementation of network lifetime is confined to the purpose of application where the sensors are deployed (Shokouhifar & Jalali, 2015). Tuning the parameters of energy consumption using a genetic algorithm and simulated annealing shows the improvement in network lifetime.

Incorporating machine learning for attaining energy efficiency with the MAC layer has been proposed (Galzarano et al., 2014). The learning agent determines the duty cycle operation where the node is either powered on or goes to sleep with its

corresponding time slot. Depending on the application a connected correlated domain set captures the correlation of sensors located within a sensing field (Gupta et al., 2008). Hyper-graph capture correlation takes place among sensors and connectivity is established with correlation edges. Thus energy efficiency has been achieved by reducing the number of transmissions.

3.2.1 Practical Implementation of Clustering Algorithm

The IEEE 802.15.4 communication through fieldbus can lead to channel outage probability and interference. Energy transceiver operation tends to be low, hence to provide consistency, functionality retransmission has to be reduced (Willig et al., 2005). Data fusion in sensor nodes has been discussed with "Tenet architecture" consisting of motes where sensor data is generated locally and gathered and processed at master or sink. Information transfer takes place within the network diameter to the master based on the task requirements (Gnawali et al., 2006). Large network diameter will result in wastage of wireless resources due redundant sensing and interference of communication. Hence small network diameter of sensor within a sensing area is preferred.

Application of wireless sensors using real time test bed discusses the necessity of time period (delivery delay of data) to perform embedded computing and optimization (Hoang et al., 2013). The optimal solution of achieving energy efficiency through clustering can be classified as software based and hardware based. In addition, the real time and non-real time reporting phenomena are also the metrics aiding the quality of service (QoS) (Amjad et al., 2017). Using clustering process cascading solutions to balance the network load has been discussed. It suggests, the cluster size is an important metrics to determine the reboot time of sensors and resist cascading failures (Fu et al., 2019).

3.3. EAUC-DUAL - Energy Aware Unequal Clustering Algorithm with Dual Sink

The proposed protocol Energy Aware Unequal Clustering algorithm with Dual sink focuses on unequal clustering and transmits data to the base station using dual sink. EAUC- DUAL algorithm uses a static sink and a mobile sink for transmission of data to the base station. “SS” denotes the static sink and “MS” denotes the mobile sink in figure 3.2. The work proposes smaller-sized clusters around the static sink and increased cluster sizes farther away from the sink. The nodes or cluster heads in the network only transmit their data to their nearest static or mobile sink. Cluster heads are elected in two phases, the primary cluster head selection (PCH) phase followed by secondary cluster head selection phase (SCH). The data is transmitted from SCH of each cluster to static or mobile sink by intra-cluster or inter-cluster routing process.

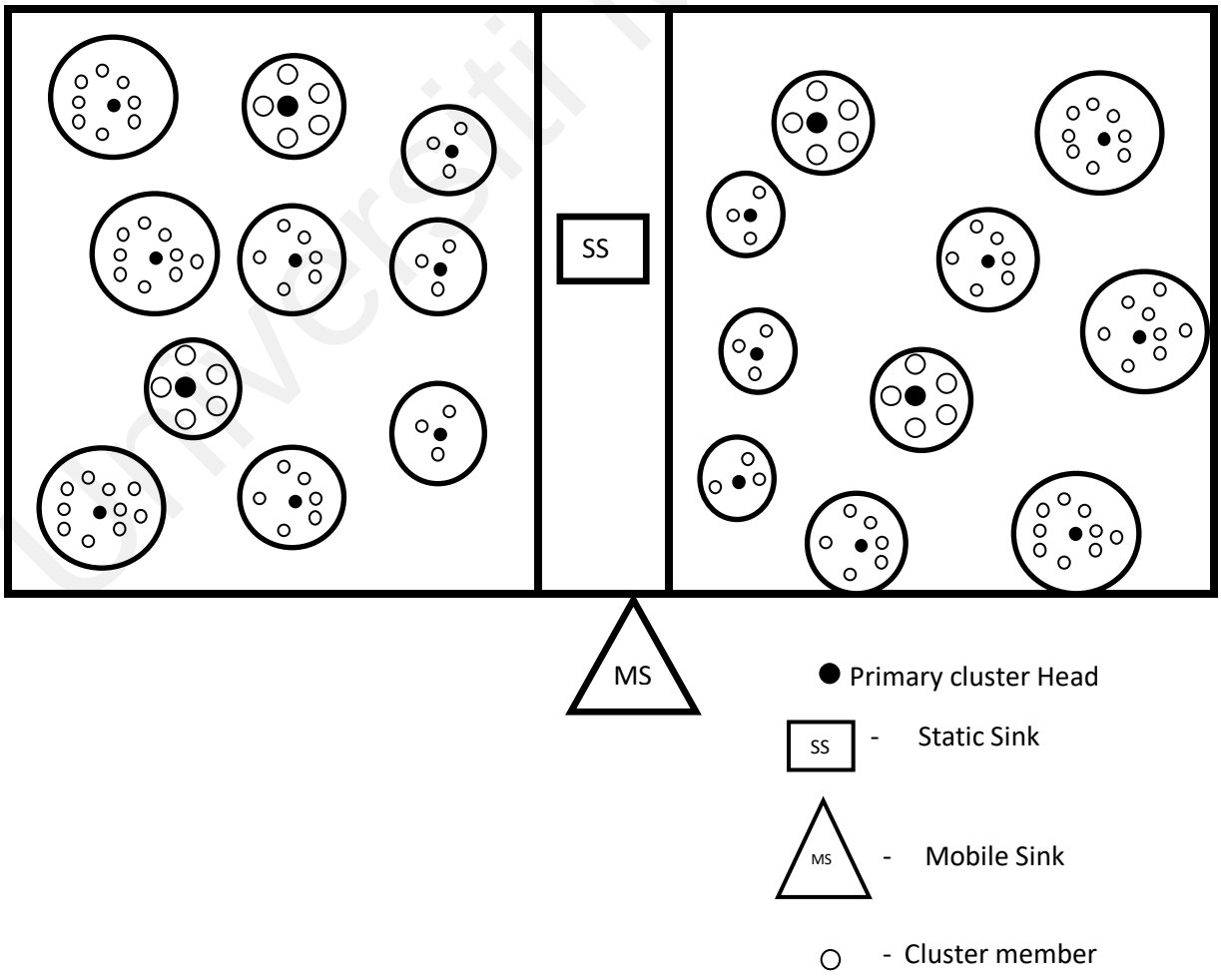

Figure 3.2: Network model of EAUC-DUAL

Each process is detailed in the following sections. Simulation results confirm, EAUC-DUAL outperforms LEACH with a static sink and LEACH with a dual sink in terms of network lifetime and energy consumption. Figure 3.2 illustrates the network model. The sensor network consists of lot of sensor nodes deployed in a square area. Unequal clustering approach is used where network is partitioned into different cluster sizes. Each cluster has a unique primary cluster head and certain number of clusters members. The following are the criteria made about the wireless sensor nods and clusters.

- There is a static sink and mobile sink located in the sensor network field. Static sink is positioned at the center of the network and mobile sink is moving in a linear pattern.
- Primary cluster heads and cluster members are stationary after deployment. The significance of the sensor network is to collect data on a continuous basis and send to the static or mobile sink by wireless way.
- Each primary cluster head and cluster member is assigned a unique identifier (ID). To recharge the batteries of sensor nodes is impossible. There is energy depletion only due to frequent transmission
- The cluster members and primary cluster head can adjust the power levels dynamically.
- Cluster members and primary cluster head do not have GPS equipment and cannot get the global position. The relative distance can be computed using the received signal strength.
- Mobile and static sink nodes are aware of the location they are currently present.

The main purpose of this work is to study the performance of dual sink strategies with wireless sensor networks transmissions in terms of energy consumption, network lifetime and also to mitigate hot spot problem. A traditional network is modelled as a graph $U(S, L)$ where S is the set of primary cluster heads and cluster members and L is the set of all links (i, j). Here i and j are neighbor nodes. The amount of energy consumed by the cluster head while aggregating data is E_{DA} (nJ/bit/signal). The radio energy model described in Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol (Heinzelman, 2000) is used in EAUC-DUAL. EAUC – DAUL protocol balances dissipation of energy among the sensor nodes and prolongs the network lifetime. The below sections present the details of EAUC – DUAL protocol.

3.3.1 Cluster Head Selection

Secondary cluster head selection is similar to the LEACH mechanism (Heinzelman, 2000). The secondary cluster heads are randomly selected using a threshold value. Each sensor node chooses a random value between 0 and 1. If the number chosen by the sensor node is less than or equal to the threshold value then the respective sensor node is selected as a secondary cluster head. Equation 3.1 defines the threshold equation.

$$T(i) = \begin{cases} \frac{SCH}{1-SCH\times\left(Round \bmod \frac{1}{p}\right)} & \text{if } i \in NSCH \\ 0 & \text{otherwise} \end{cases} \quad (3.1)$$

SCH is the desired percentage of the secondary cluster head nodes, Round is the current round number and $NSCH$ is the set of non-secondary cluster head nodes in the last $\frac{1}{SCH}$ rounds. Depending on the Round either the SCH or NSCH is chosen. This is denoted by two random values 0 and 1. Secondary cluster heads compete to become primary cluster heads depending on their competition range, residual energy, and node ID. This work focus on dual sink but the competition range is computed based only on the position of the static sink.

Each sensor node computes its distance d to the location of the static sink node and find its minimum distance d_{min} and the maximum distance d_{max} to the static sink. On this root, the sensor node calculates its competition range R_r in order to form clusters of unequal sizes. The competition range R_r is predefined in 3.2.

$$R_r = \left(1 - c\,\frac{d_{max} - d(S_i, Static\ Sink)}{d_{max} - d_{min}}\right) R_0 \qquad (3.2)$$

Here R_0 is the maximum competition range and c is a constant between 0 and 1. we can observe the competition range R_r decreases as the distance to the static sink node decreases. Each secondary cluster head has to broadcast a message including its competition range and its residual energy to its neighbour secondary cluster head. Those secondary cluster heads within the limits of the competition range R_r are defined as the neighbor secondary cluster heads. At the end of the competition only one primary cluster head is tolerable in each competition range. Figure 3.3 shows an arrangement of secondary cluster heads. The circles represent different competition ranges of secondary cluster heads. In Figure 3.3 both C_1 and C_2 could become primary cluster heads.

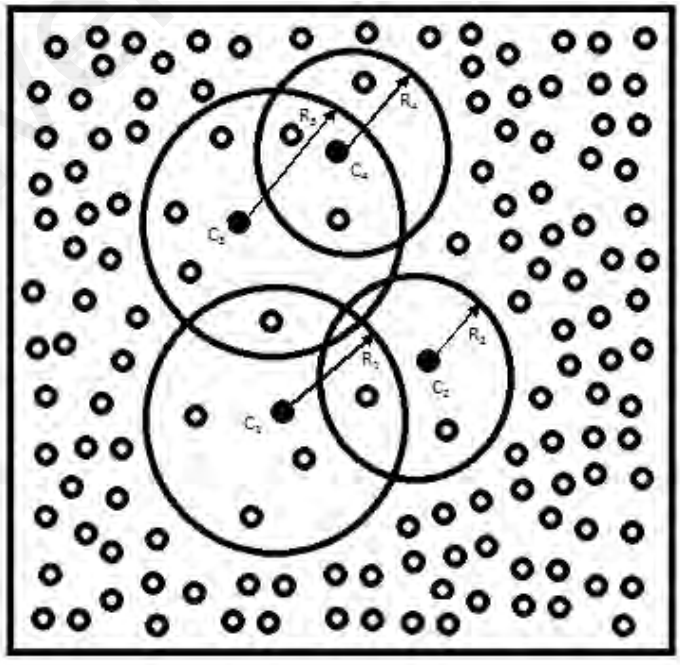

Figure 3.3: Selection of Primary Cluster Heads

The secondary cluster heads C_3 and C_4 cannot become primary cluster heads at the same time. Based on residual energy the secondary cluster head C_3 and C_4 compete with each other to become the primary cluster head. Secondary cluster head has a higher amount of residual energy and would be selected as the primary cluster head. If C_3 and C_4 secondary cluster heads have equal residual energy then the secondary cluster head with a slighter ID would be elected as the primary cluster head. The non-cluster head nodes will link itself to the nearest primary cluster heads.

3.3.2 Dual Sink Mobility Model

The usage of dual sink is supportive for continuous monitoring applications and also reduces hot spot problem. In our work we use linear path mobile sink along with static sink at the centre to collect data from primary cluster heads and cluster members. The mobile sink node moves along the linear path back and forth and gathers data packets from primary cluster head or cluster member during Intra-cluster routing or Inter-cluster routing phase. Assume x and y to be the coordinates of the initial position of mobile sink and the mobile sink node (MS) moves with a velocity of 10 metres per round. The next round position of the mobile sink node is determined in 3.3 and 3.4.

$$MS_{(x+1)} = MS_x + velocity \quad (3.3)$$

$$MS_{(y+1)} = MS_y + velocity \quad (3.4)$$

3.3.2.1 Intra-Cluster Single Hop Communication Using Dual Sink

In this proposed routing algorithm each cluster member node compares the distance to static sink $d(CM, SS)$ and mobile sink $d(CM, MS)$ along with the distance to its corresponding primary cluster head $d(CM, PCH)$. The cluster member node (CM) will send its data to its corresponding primary cluster head (PCH). In case if the distance to the mobile sink (MS) or static sink (SS) is lesser compared to the distance of its corresponding primary cluster head then the cluster member transmits its data

directly to mobile sink. The data transmission to the primary cluster head is given in equation 3.5.

$$\text{Data Transmission to PCH} = \mathrm{d}(\mathrm{CM}, \mathrm{PCH}) < d(\mathrm{CM}, \mathrm{MS}) < d(CM, SS) \quad (3.5)$$

For intra – cluster routing using dual sink each cluster member transmits data directly to the sink node (static or mobile) or to its corresponding cluster head based on its distance. Since mobile sink and static sink are deployed, in practice some cluster member may consume less energy through sending direct data to the sink than to its primary cluster head. The cluster members are generally close to their corresponding cluster heads and therefore implementation of Intra – Cluster routing could minimize delay latency.

3.3.2.2 Inter-Cluster Multi-Hop Communication Using Dual Sink

Each primary cluster head node will compare the distance to mobile sink node $d(PCH, MS)$ and distance to static sink node $d(PCH, SS)$, and then transmit the data packet to the mobile sink node only if the distance to the mobile sink node is lesser than the distance to static sink node. Otherwise the aggregated data packet will be transmitted either directly or indirectly through neighbour primary cluster head nodes $(NPCH)$ to the static sink node (SS). Equation 3.6 illustrates the data transmission to mobile sink.

$$Data\ Transmission\ to\ MS = d(PCH, MS) < d(PCH, SS) \quad (3.6)$$

Hence based on the shortest distance, each primary cluster head transmits data to the mobile sink or to the neighbour primary cluster head node. Certain primary cluster heads might be positioned far away from either sinks hence implementation of Inter-cluster routing protocol not only diminishes energy consumption but also makes this proposed algorithm advisable and relevant for large scale wireless sensor networks.

3.3.2.3 Algorithm for Inter- Cluster and Intra-Cluster Routing

Step 1:

Initialize the network setup

Step 2:

Selection of secondary cluster head
Selection of primary cluster head
Formation of unequal cluster sizes

Step 3:

Intra cluster Routing

To send the data from Cluster Member, compare distance between Primary Cluster Head, Static sink and Mobile.

$if\ distance\big((SS < PCH) and\ (SS < MS)\big)\ data\ transmitted\ to\ SS$
$if\ distance\big((MS < PCH) and (MS < SS)\big)\ data\ will\ be\ sent\ to\ MS$
$if\ distance\big((MS > PCH) and\ (SS > PCH)\big) data\ will\ be\ sent\ to\ PCH$

Step 4:

Inter cluster Routing

To send the data from Primary Cluster Head, distance between Mobile Sink and Static Sink are compared

$if\ distance\big((MS < NPCH) and\ (MS < SS)\big)\ data\ transmitted\ to\ MS\ from\ PCH$
$if\ distance\big((SS < NPCH) and\ (MS > NPCH)\big)\ data\ transmitted\ to\ SS\ from\ PCH$
$if\ distance\big((NPCH < SS) and\ (NPCH < S)\big)\ data\ transmitted\ to\ NPCH\ from\ PCH, then$
$if\ distance\big((NPCH < SS) and\ (NPCH > MS)\big)\ data\ transmitted\ to\ SS\ from\ NPCH$

Repeat the above step until data packet is transmitted to the sink

Step 5:

Next round, repeat step 2.

LEACH protocol selects cluster head in a random manner leading to maximum energy availability at the specific instant of time. The protocol assumption is made in such a way where all nodes can transfer data to sink. However, the long detour paths involved in data transfer of alternate cluster head consumes more energy and decreases the network lifetime. In EAUC-DUAL the data transfer takes place considering computational capability of sensor nodes in choosing the primary cluster heads and also based on distance to sink.

3.4 Results and Discussion of EAUC-DUAL

EAUC-DUAL for continuous monitoring of applications in wireless sensor networks combines an unequal clustering algorithm with dual sink in order to improve network performance by reducing energy consumption, prolonging network lifetime and mitigating hotspot problems. The operation of the proposed routing algorithm is divided into rounds. There are 100 sensor nodes dispersed in a 300m × 300m homogeneous wireless sensor network with a static sink node placed in the center of the network as well as a mobile sink node moving along the linear path back and forth at 10 m per round. Cluster member nodes transmit a 4,000-bit data packet using an intra-cluster routing protocol to either their corresponding PCH nodes or the sink node depending on the minimum distance. PCH nodes aggregate data packets from cluster members and transmit them using the inter-cluster routing protocol to either the static or the mobile sink node depending on the minimum distance. The performance of Dual Sink based - Energy Aware Unequal Clustering Routing Algorithm (EAUC-DUAL) is evaluated using MATLAB simulations. Simulation parameters are specified in table 3.1.

Table 3.1: Simulation parameters of clustering algorithms

Parameter	Values
Network Field	300 m×300 m
Number of Sensor Nodes, N_s	100
Initial Energy, E_0	0.5 Joules
Electronics Energy, E_{elec}	50 nJ/bit
Data Aggregation Energy, EDA	5 nJ/bit
Transmit amplifier free space propagation model, $\in_{fs}$	10 pJ/bit/m^2
Transmit amplifier two-ray ground propagation space model, ε_{mp}	0.0013 pJ/bit/m^4
Percentage of Secondary cluster heads, SCH	0.2
Maximum competition range, R_0	40 m
Data Packet Size	4000 bits

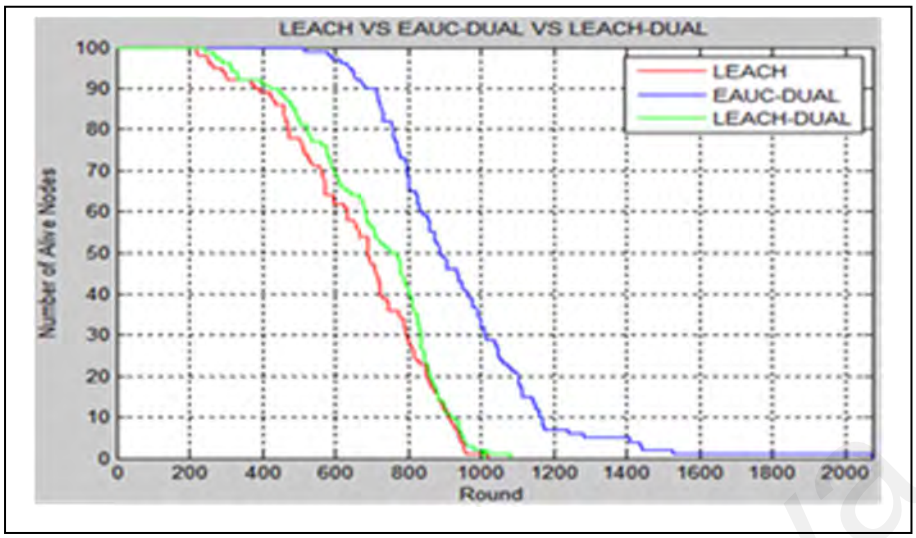

Figure 3.4: Number of Alive Nodes

Table 3.2: Comparison of Network Lifetime

Algorithms	**Network Lifetime (Rounds)**
LEACH	220
LEACH using Both Static Sink and Mobile Sink (Linear Path Mobility Pattern) (LEACH-DUAL)	242
Energy Aware Unequal Clustering Routing Algorithm using Both Static Sink and Mobile Sink (Linear Path Mobility Pattern) (EAUC-DUAL)	513

The performance of the proposed clustering routing algorithm is compared with LEACH using static sink and LEACH using both static sink and mobile sink (LEACH – DUAL). Simulation results are shown in figure 3.4. From the graph it is clear EAUC – DUAL has increased number of rounds than LEACH and LEACH-DUAL. Table 3.2 shows the network lifetime comparison between LEACH, LEACH- DUAL and EAUC-DUAL. We define the network lifetime as the time period until the first node depletes its own energy. As seen from table 3.2, there is an obvious increase in network lifetime of the proposed clustering routing algorithm as compared to other routing algorithms. The network lifetime of the proposed clustering routing algorithm (513 rounds) is longer than LEACH (220 rounds) and LEACH using both static sink and mobile sink (242 rounds).

3.5 ECH-DUAL - An Energy based Cluster Head Selection Unequal Clustering Algorithm with Dual Sink.

In the previous algorithm EAUC - DUAL the sensing field is small of 300m × 300 m and hence clustering and transfer of data to sink has been effective. An ECH-DUAL algorithm is proposed in this section where the sensing field is increased and further cluster head selection method is modified to enhance the energy consumption. In the scenario of ECH DUAL, sensing field is increased to 1000 m ×1000 m hence the distance between sensors increases and simultaneously the energy consumption of nodes also increases. So to counter the issues of energy consumption a novel procedure for election of tentative cluster head followed by final cluster head is proposed. In addition to static sink in the network, mobile sink also navigates with awareness to alleviate hotspots. Figure 3.5 illustrates the overall flow of ECH-DUAL algorithm. The network is initialized.

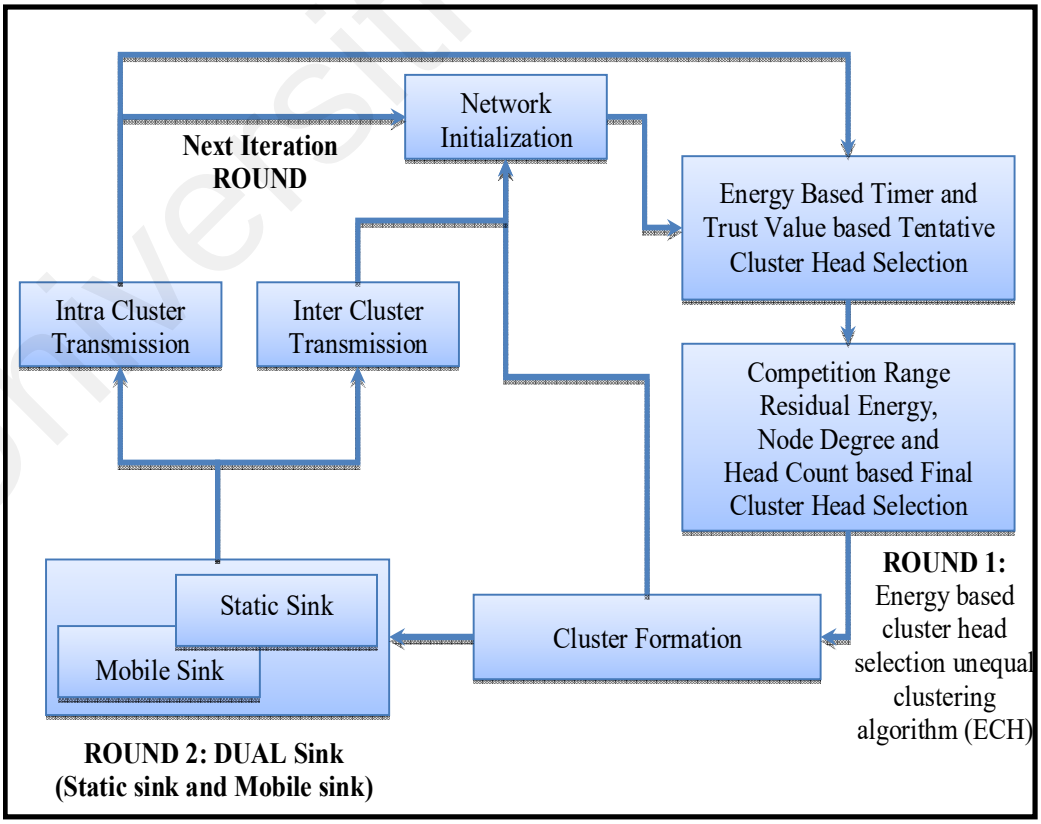

Figure 3.5: Illustration of Energy based Cluster Head Selection Unequal Clustering with DUAL sink (ECH-DUAL)

The network lifetime of sensor nodes from the initialization to the transfer of data is determined by appropriate clustering of sensor nodes. The clustering and cluster head selection criteria based on energy metrics increases the network lifetime of nodes. Cluster Head selection is one of the crucial parameters balancing both energy and delay parameters when associating to its cluster members (Thakkar & Kotecha, 2006). Hence selection of cluster heads cannot be fixed and tends to be varying. A novel cluster head selection process is explained involving two stages namely tentative cluster head selection and final cluster head selection. In this algorithm tentative cluster heads are selected based on the energy-based timer (EBT) and the trust value (TV). The final cluster head selection approach is based on competition range, sensor energy and node ID. Thus unequal size clusters are formed. Data transmission takes place using dual sink. One of the sinks is positioned at the center of the topography while the other sink moves in a linear path back and forth in a linear pattern. The sinks collect data from final cluster head through inter-cluster and intra-cluster routing.

Figure 3.6 represents the network model of ECH-DUAL. This model consists of many clusters of unequal sizes. The clusters near the sink is smaller than those clusters farther away so as to reduce the hot spot problem. The final cluster head is formed in each round depending on the signal strength and radio range. The static and mobile sinks are located at networks center path and its function is to gather the data from the final cluster heads. In the network center the static sink is fixed and mobile sink moves in a straight line on the network region. The nodes are clustered in each round and the data is collected by static and mobile sink. If the static sink is nearer to the CH the collected data will be transferred to the static sink. Meanwhile if the mobile sink is nearer to the CH the collected data will be transferred to the mobile sink.

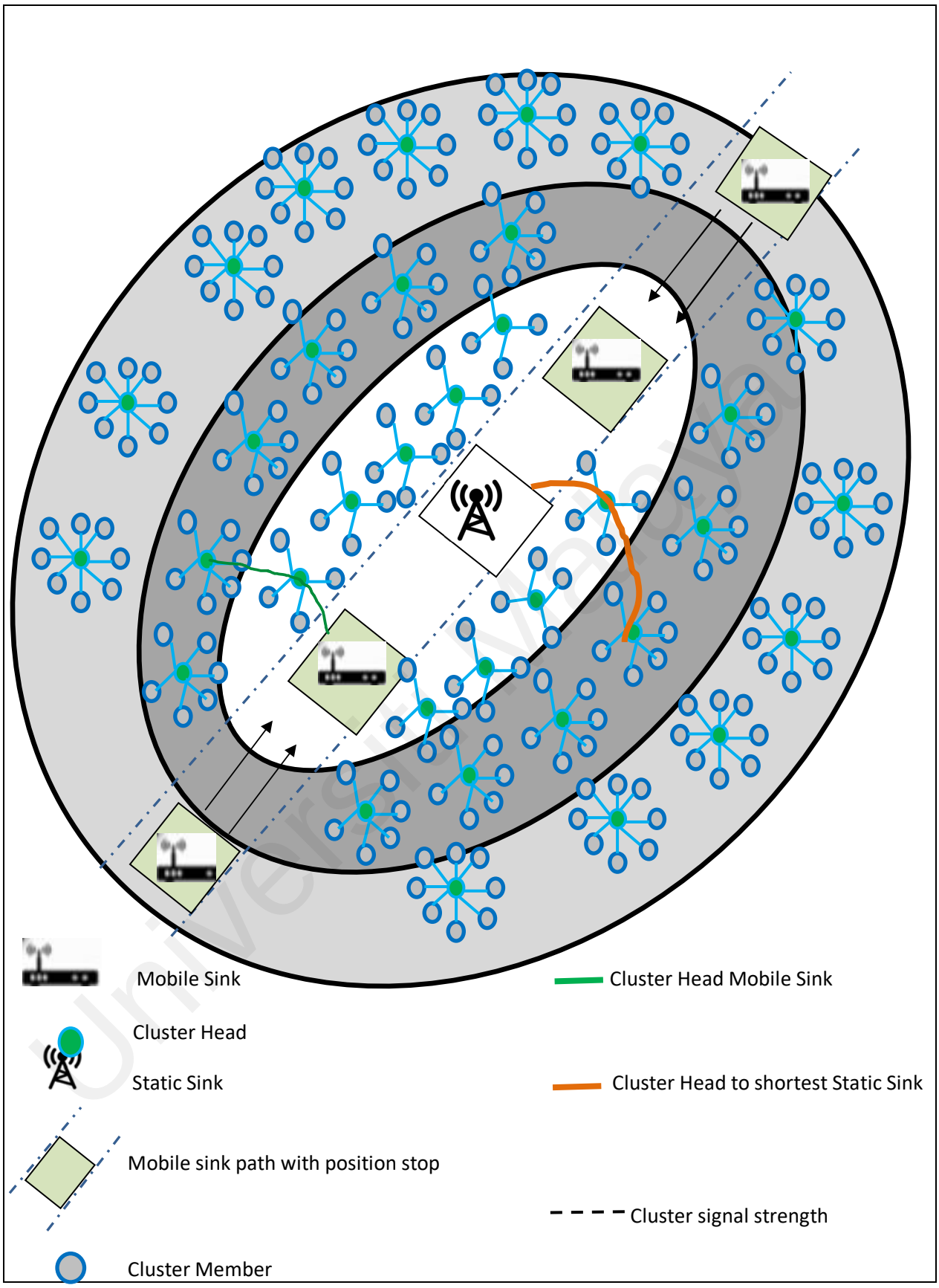

Figure 3.6: Network model of ECH-DUAL

The same technique is applied to all clusters for effective communication thereby reducing the communication cost and increased network lifetime. Thus, the proposed **ECH – DUAL** algorithm emphasis on cluster head selection and routes the data from the final cluster head to static or mobile sink. The protocol overcomes the funneling effect and hotspot problems of sensor nodes by means of dual sink and also by incorporating unequal clustering using novel cluster head selection method.

3.5.1 Tentative Cluster Head Selection

Cluster head selection is based on two processes namely: Temporary Cluster Head (TCH) selection and Final Cluster Head (FCH) selection. Cluster Head is selected using tentative cluster head selection process based on Energy Based Timer (EBT) and Trust Value (TV). The timer is assigned to the node to choose the TCH and trust values are computed based on node's overall Trust value. Node possessing highest trust value and Energy is chosen as TCH. In addition to this final cluster head selection is based on competition range, node degree and head count. The sensor nodes are assigned to timer based on each node's energy. The waiting time assigned to the nodes is based on energy. The waiting time is assigned by using two criteria: the nodes with higher energy will be assigned lesser waiting time and the nodes with lower energy will be assigned higher waiting time.

The node whose timer value expires first would be selected as the Tentative Cluster Head (TCH). This process promotes high energy nodes as the next tentative cluster head. Otherwise the same node of highest transmission energy acts as the cluster head. This energy-based timer has the following model description. Suppose for a node i there are k neighbor nodes and each node can calculate the average energy value of their neighboring nodes: $S_i = \{i_1, i_2, i_3, \ldots i_n \ldots . i_k, \}$ and i_n denotes the nth neighbor node. The following equation gives the average energy of node i:

$$Average\ Energy(i) = \begin{cases} \frac{1}{k}\sum_{n=1}^{k} Energy(i_n) & k > 0 \\ 0 & k = 0 \end{cases} \qquad (3.7)$$

TCH is selected from the sensor nodes using the Energy based timer. For any sensor node ID S_i, energy-based waiting time value can be obtained from the equation 3.8. The waiting time decreases as the energy of the node increases. The node with higher energy will be assigned less waiting time. This node is selected as the tentative Cluster Head. The selected tentative CH broadcast tentative CH message in its broadcast range and other sensor nodes exits the cluster head selection upon receiving this message before the arrival of its waiting time.

$$WaitTime\ (s_i) = \frac{Average\ Energy\ of s_i Neighbornode}{Energy of S_i} \qquad (3.8)$$

Trust Value (TV) is used to detect the node behavior, node quality and node services. It is also used for data aggregation, reconfiguring and routing of sensor nodes. It provides a quantitative way to evaluate the trustworthiness of sensor nodes (Rajaram et al., 2014). Trust value is used to collect data and monitor different events in the node. Along with Energy based Timer (EBT) the trust value is used to find the tentative cluster head. Tentative CH selection follows two approaches (EBT and TV) to optimize best cluster head selection efficiency. The following equation 3.9 is used to calculate the trust value of nodes.

$$Trust\ Value(TV)_{nodes} = \frac{N_{FD}}{N_{REC}} \qquad (3.9)$$

Where N_{FD} denotes number of forwarded packets and N_{REC} denotes number of received packets. The trust values of the individual node are computed and the node with highest trust value is selected as temporary cluster head. Finally, the EBT and the TV returns the result of TCH selection. After this process, final CH selection is performed.

3.5.2 Final Cluster Head Selection

TCH compete to become FCH based on parameters such as competition range, residual energy, node degree and head count. Based on the sensor node's energy consumption dead and alive nodes of the cluster are identified. Final Cluster head is selected based on the following process. Number of edges incident on a node is called the node degree (Bettstetter, 2002).

The average node degree of N is denoted by:

$$d_{mean}(N) = \frac{1}{n}\sum_{n=1}^{N} d(n) \qquad (3.10)$$

Where d (n) is the degree of a node, 'n' is the number of neighbors of node (links). A node of degree 0 has no neighbors. The minimum node degree of a network 'N' is defined as:

$$d_{min}(N) = \min_{\forall n \in N}\{d(n)\} \qquad (3.11)$$

The d_{min} and d_{mean} gives efficient results compared to EAUC with dual sink. Node whose degree is higher is elected as CH. Node degree reduces the overall communication cost for cluster-head selection and thus it increases overall lifetime of the network. Each normal node belongs to only one cluster. Minimizing number of clusters maximizes average cluster size. The advance nodes deployed in dense areas are selected as cluster head. The competition range among tentative cluster heads is shown in figure 3.7. With respect to the chosen TCH, FCH are selected using following process. Consider S1, S2, S3, S4, S5 are selected as five tentative cluster heads. Each TCH will compute the competition range R_i, using the following formulae:

Case 1: Competition range for tentative cluster head (S1)

$$S1(R)_1 = \left(1 - c\,\frac{d_{max} - d(S_1, Static\ Sink)}{d_{max} - d_{min}}\right) R_0 \qquad (3.12)$$

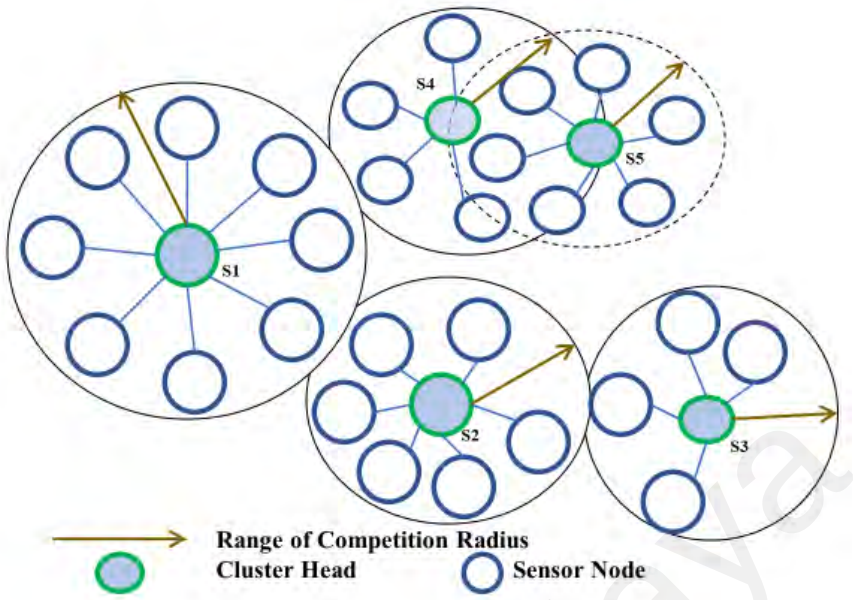

Figure 3.7: Competition range of Tentative Cluster Heads

Case 2: Competition range for tentative cluster head (S2)

$$S2(R)_2 = \left(1 - c\frac{d_{max} - d(S_2, Static\ Sink)}{d_{max} - d_{min}}\right) R_0 \qquad (3.13)$$

Case 3: Competition range for tentative cluster head (S3)

$$S3(R)_3 = \left(1 - c\frac{d_{max} - d(S_3, Static\ Sink)}{d_{max} - d_{min}}\right) R_0 \qquad (3.14)$$

Case 4: Competition range for tentative cluster head (S4)

$$S4(R)_4 = \left(1 - c\frac{d_{max} - d(S_4, Static\ Sink)}{d_{max} - d_{min}}\right) R_0 \qquad (3.15)$$

Case 5: Competition range for tentative cluster head (S5)

$$S5(R)_5 = \left(1 - c\frac{d_{max} - d(S_5, Static\ Sink)}{d_{max} - d_{min}}\right) R_0 \qquad (3.16)$$

As the distance $d(S_i, \text{Static Sink})$ increases the competition range R_i of the tentative cluster head also increases and vice versa. For example, S3 is nearer to the static sink and hence the distance between static sink and S3, {d (S3, static sink)} is small. Hence the competition radius $S3(R)_3$ will also be small. Thus the communication range of S3 will be small. Hence the sensor nodes in the respective small range join S3 creating a small cluster, with S3 as the final cluster head as there is no other tentative

CH overlapping in the same range. Similarly S1 and S2 calculate the competition range. The range is decided based on the distance to the sink. As S1 and S2 are far from sink their competition range should be larger than the S3. They do not have competition or overlap as no other tentative CH is present in their range. Hence S1 and S2 become FCH. The nodes in the competition range of S1 and S2 join them as cluster members. In another case the TCH's S4 and S5 overlap in a cluster. S4 and S5 compete to become FCH. S4 can hear the broadcast message of S5 and vice versa as their ranges overlap. The competition between S4 and S5 to become Final CH is explained below:

- If S5 belongs to S4($S5\ \epsilon S4$). S4 compares its energy with the energy of S5
- E (S4> $S5$) then S4 broadcast itself as final CH. S4 receives a quit election message from S5. S4 removes S5 from its overlap region. Thus S4 will become the FCH for the cluster in transmission range. Other sensor nodes in S4 range will join S4 as its cluster members
- E (S5>S4) then S5 sends FCH message to S4. After S4 receives the message from S5 it gives up and sends a quit election message to the nodes in overlap region. Hence S5 will be the final CH for the cluster in its range. All nodes in S5 range will join S5 as its cluster members.
- If the energy of S4 and S5 are equal (tie) then the node with smallest id will become the final cluster head.

The pseudo-code for CH selection in unequal clustering algorithm provides an efficient energy balance in the network. Figure 3.8 explains the Pseudo code for tentative CH selection and final CH selection. The Energy Based Timer (EBT) and trust value are used to choose TCH. The FCH is selected based on the node degree, competition range, residual energy and head count.

Pseudo-code for Tentative CH and Final Cluster Head Selection

Initialization:

Energy (E), Temp_CH(Sj), Total Energy (E_{total}), Forwarded packets (N_{FD}), Total Number of Packets (N_{REC}), non-cluster set of nodes (G), Number of Node (N)

Case 1: *Tentative CH Selection ∀ node (i)∈ [Within Network Region]*

energy(i)←Energy level of node i

initialize timer (i)←'k', neighbor node, i_n n^{th} neighbor node with sensor node S_i

elect cluster head as 'n'

if n become the cluster head

process checks the cluster distance of

d_{max}= max (d_{max}, distance) and d_{min}=min(d_{min}, distance)

Calculate the average energy

$$Avg\ Energy\ (i) = \begin{cases} \frac{1}{k}\sum_{n=1}^{k} Energy(i_n) & k > 0 \\ 0 & k = 0 \end{cases}$$

until end of the node in *non-cluster set of nodes (G)*

Wait Time (S_i)= (Average Energy of S_i Neighbor node)/(Energy of S_i)

repeat process to reach successive tentative Cluster Heads with high probability

Case 2: *depends on the Random node (n) and non-cluster set of nodes (G)*

if 'n' sends advertisement and join request to G

Non-cluster set of nodes *checks and verifies* the cluster head

d_{max} and $d_{min}$with the total energy E_{total}=E_{total}+ E

else

Create cluster head (CH) for new cluster

form the remaining *non-cluster set of nodes*

perform case 1: and case 2:

calculate the trust value, $(TV)_{nodes}$=N_{FD}/N_{REC}

Result: tentative cluster head selection

Case 3: *find final cluster head selection*

compute the node degree by

$d_{min} = \min_{\forall n \in N}\{d(n)\}$, average node degree by

$d_{mean}\ (N) = \frac{1}{n}\sum_{n=1}^{N} d(n)$

for all dense area

Calculate competition range of tentative cluster head by

$$R_i = \left(1 - c\frac{d_{max} - d(S_i, \text{Static Sink})}{d_{max} - d_{min}}\right) R_0$$

If distance $d(S_i, \text{Static Sink})$ is increases

Competition range R_i increases

else if distance $d(S_i, \text{Static Sink})$ is decrease

Competition range R_i decrease

Condition 1: cluster size w.r.t. sink by near → small

Condition 2: cluster size w.r.t. sink by far → large

Select the CH as FCH;

Result:FCH selection

Figure 3.8: Pseudo code for Tentative CH and Final CH selection

3.5.3 Energy Consumption

The Residual Energy (RE) of proposed algorithm is measured in each round starting from the current Round, Round+1 and Round+ 2 until final node is reached. The energy consumption rate of proposed system is shown in equation 3.17 to 3.20. The energy consumption measured in various rounds is as follows:

In current round the energy consumed by CH is given in equation 3.17.

$$ResidualEnergy = RE + S(i) * E \quad (3.17)$$

From this the average residual energy (ARE) is calculated using,

$$ARE(Round + 1) = \frac{RE(Round+1)}{2} \quad (3.18)$$

The Total Energy Consumption (TEC) on each round are calculated using the following equation.

$$TEC(Round + 1) = E_0 * n - RE(Round + 1) \quad (3.19)$$

In case there are 'n' layers available in the network the average energy consumption of node can be defined as:

$$AEC(Round + 1) = \frac{TEC(Round+1)}{n} \quad (3.20)$$

The AEC is calculated with respect to the total energy consumption. TEC consists of the average of all transmitted energy, received energy, idle energy and sleep mode energy.

3.5.4 Static and Mobile Sink in Network Model

The proposed dual sink used for aggregated data transmission network model is based on the static and mobile sink. At the center of the network the static sink is fixed to collect data from nearby CH and mobile sink moves in a linear path on the network region towards the static sink. The CHs gather data from their cluster members. This data can be sent directly or through intermediate CHs to the static sink. As a result, the static path is created and hot spot problem may occur due to energy dissipation in WSN. Hence mobile sink is also used to collect data so as to avoid hotspot problem. The coordinates of static sink are given by equation 3.21 and 3.22. The mobile sink travels

along a straight path across the entire network preventing premature death of sensor nodes. The coordinates of mobile sink are given by equation 3.23 and 3.24.

$$Sink.Xfixed = 0.5 * Xm \quad (3.21)$$

$$Sink.Yfixed = 0.5 * Ym \quad (3.22)$$

$$Sink.Xmobile = 0.5 * Xm \quad (3.23)$$

$$Sink.Ymobile = 0 * Ym \quad (3.24)$$

In equation 3.21 X_m denotes the x coordinates of the sink in topography. In equation 3.22 Y_m denotes the y coordinates of the sink in topography. The figure 3.9 depicts the flowchart for inter cluster communication where the cluster head decides to communicate to static or mobile sink based on the distance.

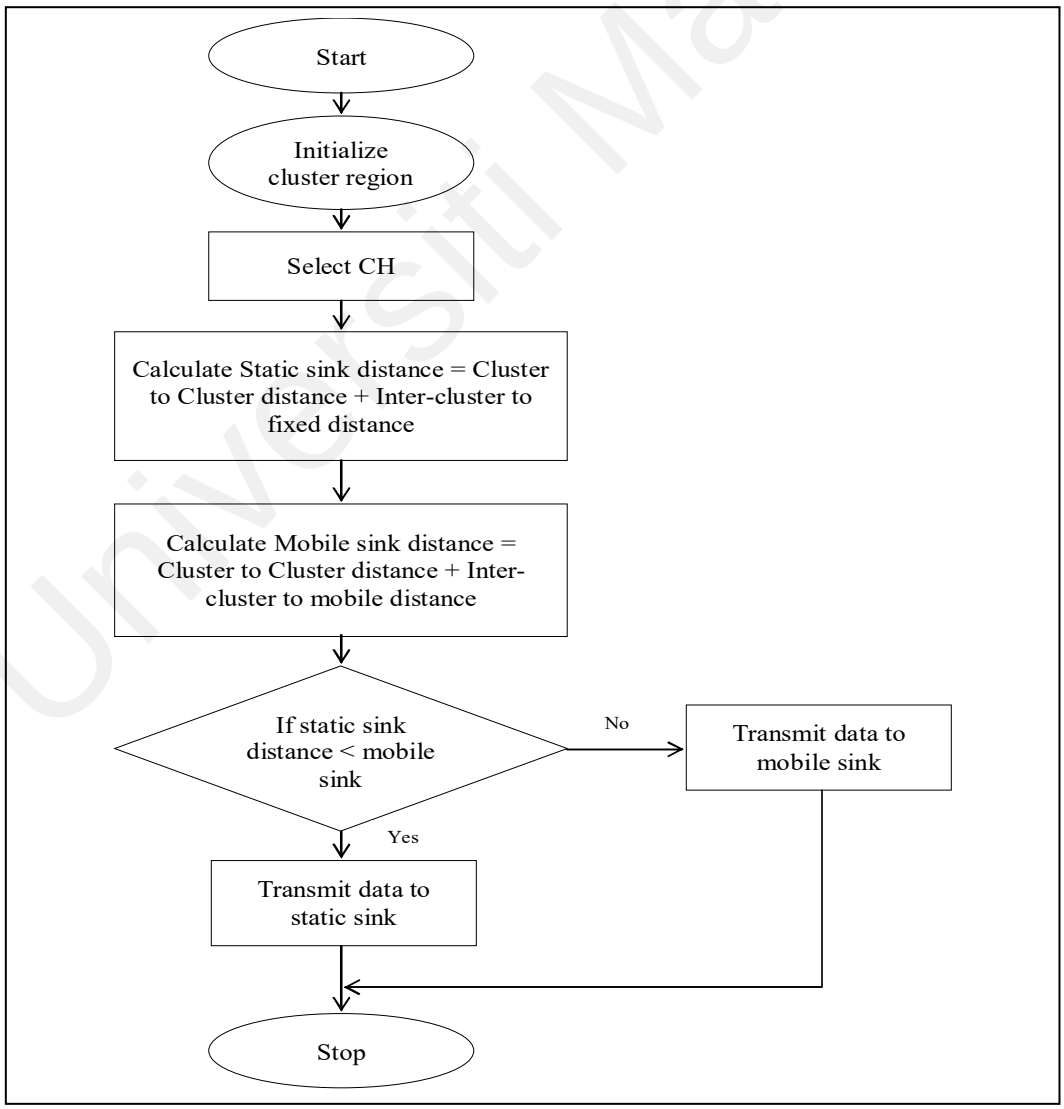

Figure 3.9: Flow chart for Inter cluster communication

3.6 Results and Discussion of ECH-DUAL

The characteristics of each node in the network and its performance are analyzed based on efficient cluster head selection and data transmission using dual sink. The proposed methodology is tested using MATLAB. In this work the terrain area of 1000 m ×1000 m is simulated with 100 nodes. The Maximum competition range (R_0) 40m is assumed as the network region.

The proposed ECH –DUAL is compared with EAUC algorithm. In addition the experiments are conducted to determine the number of alive nodes, dead nodes, residual energy, average residual energy, total energy consumption, average energy consumption and the network lifetime. The simulation iteration rounds vary from round 1 to 4250. The simulation is carried out by using the parameters given in table 3.3.

Table 3.3: Simulation parameters of ECH-DUAL and EAUC-DUAL

Parameters	Values
Network Field	1000 m ×1000 m
X dimension of topography	1000m
Y dimension of topography	1000 m
Number of Sensor Nodes, N	100
Initial Energy, E_0	0.5 Joules
Electronics Energy, E_{elec}	50 nJ/bit
Data Aggregation Energy, EDA	5 nJ/bit
Transmit amplifier two-ray ground propagation space model, φ_{mp}	0.0013 pJ/bit/m^4
Percentage of tentative cluster heads, PCH	0.2
Maximum competition range, R_0	40 m
Data Packet Size	4000 bits
Antenna Model	Omni Antenna
Simulation time	200s (Minimum:200s, Maximum:10000s)

EEUC - Energy Efficient Unequal Clustering (Li et al., 2005) is used for comparisons. In EEUC protocol unequal clustering occurs based on competition range. range. The size of clusters increases monotonically for nodes further away from static sink. The work uses two sinks (static and mobile). Euler approach has been used in static sink (data is collected at a point in space at the same instance of time) and Lagrangian approach has been used in mobile sink. The outcome produces less energy consumption and increases network lifetime and also avoids hot spot problem. At the time of simulation each node has its own power or energy. The energy is used to communicate with the neighbor node for data transmission. This type of nodes is called as the Alive Nodes. Alive nodes are identified based on balanced energy distribution and nodes lifetime. When the energy of a node become empty the node becomes dead. The alive nodes with 4250 randomly selected rounds are shown in figure 3.10. The proposed ECH-DUAL is compared with the EAUC-DUAL. The proposed algorithm maintains the alive node count high until the system reaches 4250 rounds.

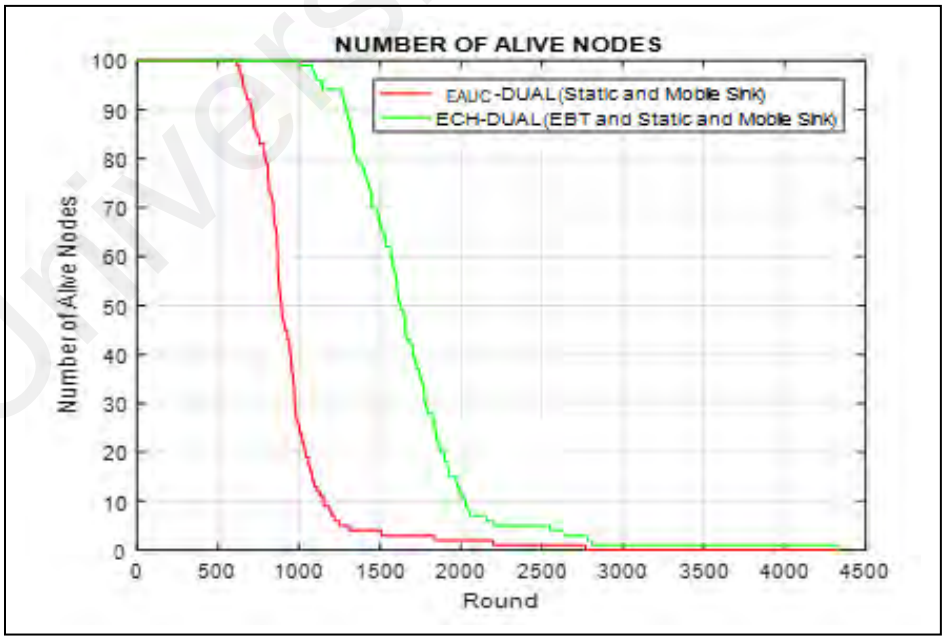

Figure 3.10: Number of Alive Nodes versus Rounds

Table 3.4: Comparison on Number of Alive nodes.

Rounds	Number of Alive Nodes ECH – DUAL (Proposed)	Number of Alive Nodes EAUC– DUAL
250	100 nodes are alive	100 nodes are alive
500	100 nodes are alive	100
1000th	100 nodes are alive	28
1120th	99 nodes alive (1st node dies)	6
1500	70 nodes are alive	4
2000	12 nodes are alive	3
2500	5 nodes are alive	1
4250	0 nodes	0

Figure 3.11 reveals the number of dead nodes for different number of rounds and the dead node count is low for proposed system. The first node dies at 1120 rounds and it is related to the network lifetime. Table 3.4 shows the comparison on number of alive nodes at different rounds.

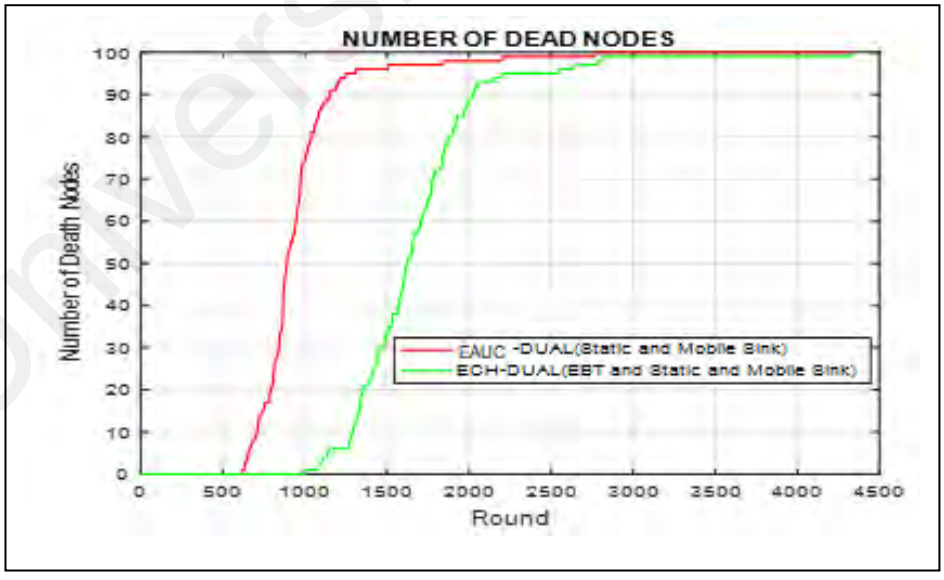

Figure 3.11: Number of Dead Nodes versus Rounds

Table 3.5: Comparison of Number of Dead Nodes

Rounds	Number of Dead Nodes ECH DUAL – (Proposed)	Number of Dead Nodes EAUC- DUAL
250	0 nodes	0
500	0 nodes	0
1000th	0 nodes	72
1120th	1 node dies	94
1500	30 nodes die	96
2000	88 nodes die	97
2500	95 nodes die	99
4250	100 - All nodes die	100

Table 3.5 depicts the number of dead nodes at different rounds. Figure 3.12 provides the average residual energy of each node with 4250 randomly selected rounds. The proposed ECH-DUAL is compared with EAUC-DUAL and the proposed system maintains less average energy consumption compared to the existing system.

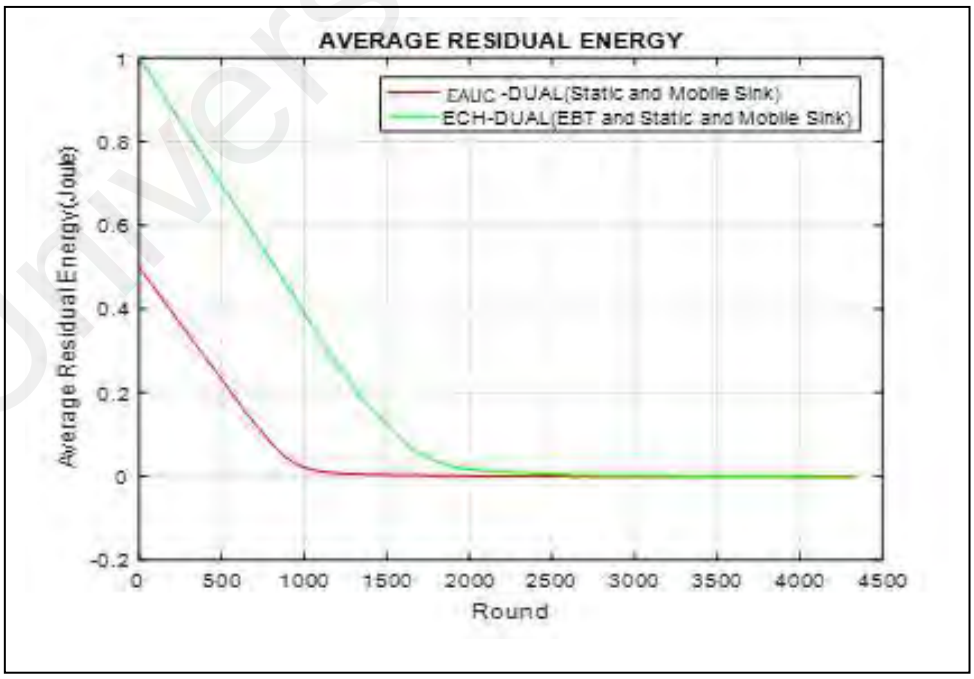

Figure 3.12: Average Residual Energy with respect to Rounds

Table 3.6: Comparison on Average Residual Energy

Rounds	Average Residual Energy – ECH DUAL (Proposed)	Average Residual Energy – EAUC DUAL
250	0.9 Joules	0.37 Joules
500	0.7 Joules	0.22 Joules
1000^{th}	0.4 Joules	0.02 Joules
1120^{th}	0.37 Joules	0.01 Joules
1500	0.10 Joules	0 Joules
2000	0.04 Joules	0 Joules
2500	0.01 Joule	0 Joules
4250	0 Joule	0 Joules

Table 3.7: Comparison on Network Lifetime

Algorithms	Network Lifetime (Rounds)
EAUC-DUAL	513
PROPOSED - ECH-DUAL	1120

Table 3.6 shows the average residual energy of sensors nodes obtained at different rounds. The Network lifetime for EAUC-DUAL and ECH-DUAL is given in table 3.7 and it is observed ECH-DUAL method returns higher network lifetime. Figure 3.13 shows comparison of lifetime of network (energy consumption) for the two algorithms with fixed dimension of 1000m ×1000m. From figure 3.13 the proposed ECH-DUAL achieves high network lifetime when compared to EAUC-DUAL.

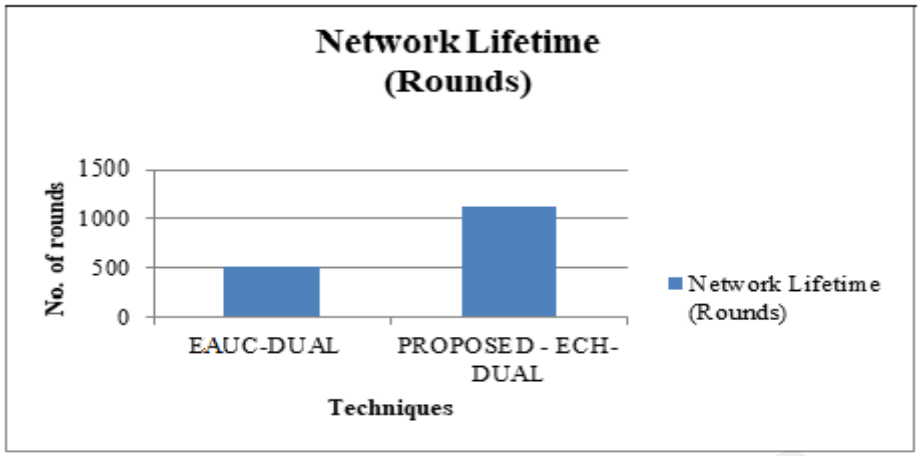

Figure 3.13: Comparison of Network Lifetime in terms of Rounds

3.6.1 Scalability of Nodes

In CAST WSN (Baradaran & Navi, 2017) initially the cluster density is chosen as 3 and the cluster size is gradually increased to 10 based on the distance coordinates. New position coordinates are being determined by using the energy center and cluster heads. Quality index of a cluster decides whether to reinitiate clustering or not. In CAST WSN two sinks are positioned and they are static in nature. ECH-DUAL positions a static and mobile sink. Table 3.8 and table 3.9 shows the simulation parameters of ECH-DUAL and CAST WSN with increased number of nodes and increased network area.

Table 3.8: Simulation parameters of CAST WSN

Protocol Name	Number of nodes	Initial energy		Total energy		Number of clusters	Monitoring area
		Minimal	Maximal	Minimal	Maximal		
CAST WSN	500	0.4	0.7	200	350	10	1000 m× 1000 m

Table 3.9: Simulation parameters of ECH-DUAL

Protocol Name	Number of nodes	Initial energy	Total energy	Sink	Monitoring area
		0.5 J	250	One static sink	
ECH DUAL	500			One mobile sink	1000 m × 1000m

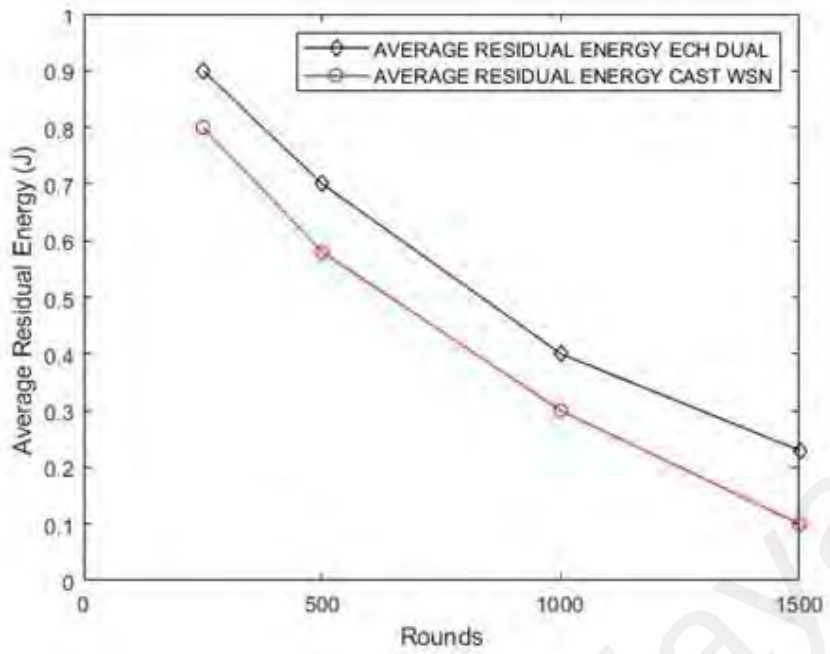

Figure 3.14: Average Residual Energy versus Rounds

Figure 3.14 and table 3.10 shows the average residual energy with respect to rounds for ECH-DUAL and CAST WSN. The average residual energy is higher in the case of ECH-DUAL due to rotation of cluster heads and mobility of one sink. The rotation and selection of Final Cluster Head in ECH-DUAL comprises of various metrics as discussed in the above section. The energy consumption is increased in CAST WSN due to static center and cluster head selection process is based only on the distance metrics. The simulation parameters of ECH-DUAL and EAUC-DUAL are shown in table 3.11. The network area is increased to 2000 m×2000 m. Also 1000 sensor nodes are used for simulation in the network area.

Table 3.10: Comparison of Average Residual Energy ECH-DUAL and CAST WSN

Rounds	Average Residual Energy ECH-Dual	Average Residual Energy CAST WSN
250	0.9	0.8
500	0.7	0.6
1000	0.4	0.3
1250	0.2	0.1

Table 3.11: Simulation parameters – Increased network area

Parameter	Values
Network Field	2000 m×2000 m
Number of Sensor Nodes, N_s	1000
Initial Energy, E_0	1 Joules
Electronics Energy, E_{elec}	50 nJ/bit
Data Aggregation Energy, EDA	5 nJ/bit
Transmit amplifier free space propagation model, ϵ_{fs}	10 pJ/bit/m^2
Transmit amplifier two-ray ground propagation space model, ε_{mp}	0.0013 pJ/bit/m^4
Percentage of Secondary cluster heads, SCH	0.2
Maximum competition range, R_0	40 m
Data Packet Size	1000 bits to 4000 bits

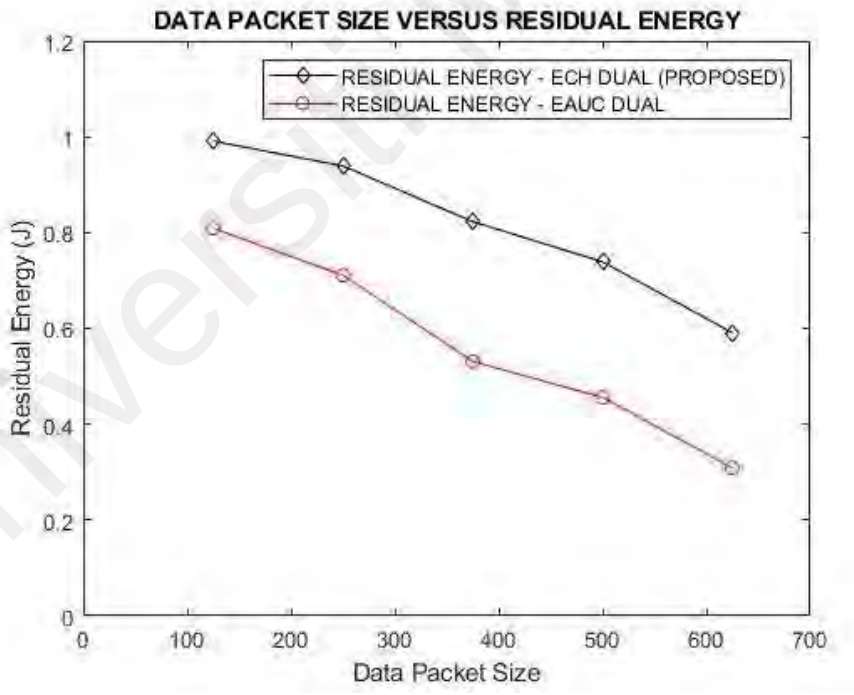

Figure 3.15: Residual Energy versus Data packet size.

In figure 3.15 the residual energy of sensor nodes is measured by gradually increasing the packet size. Residual Energy decays in both protocols with increasing packet size. ECH-DUAL achieves better energy utilization with clustering compared to conventional EAUC-DUAL.

Table 3.12: Comparison on Data Packet Size and Residual Energy

Data Packet Size	Residual Energy ECH DUAL (Proposed)	Residual Energy – EAUC DUAL
125 B	0.991 J	0.809 J
250 B	0.939 J	0.711 J
375 B	0.823 J	0.531 J
500 B	0.739 J	0.456 J
625 B	0.591 J	0.309 J

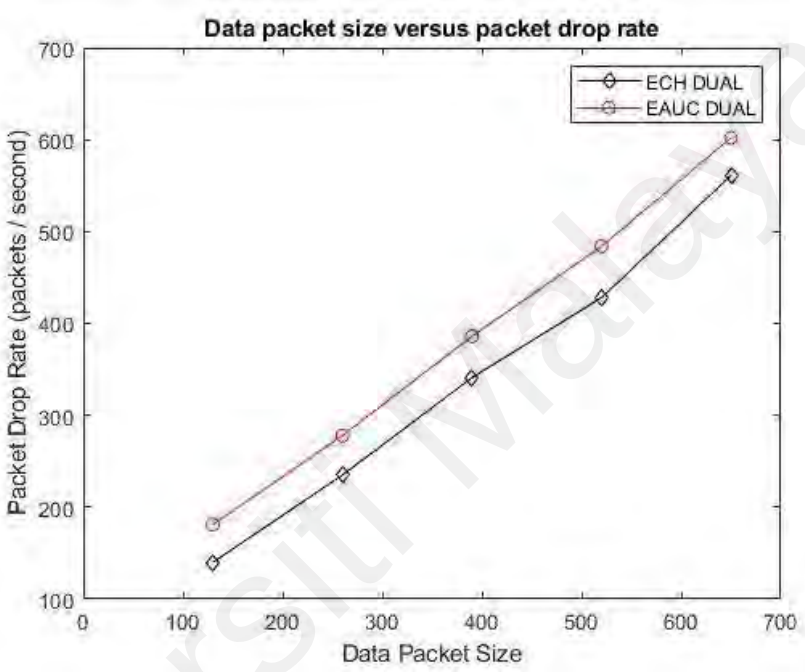

Figure 3.16: Packet Drop Rate versus Data Packet Size

Table 3.12 shows the comparison values on data packet size and residual energy. Trade- off between Data packet size and packet drop rate is shown in table 3.13 and figure 3.16. Both the protocols exhibit increased drop rate due to faulty links, hence are unable to route the packets. ECH-DUAL has reduced drop rate when compared to conventional EAUC-DUAL.

Table 3.13: Comparison on Packet Drop Rate

Data Packet Size	Packet Drop rate-ECH DUAL (packets/second)	Packet Drop rate-EAUC DUAL (packets/second)
125 B	140	181
250 B	235	278
375 B	341	386
500 B	428	484
625 B	561	602

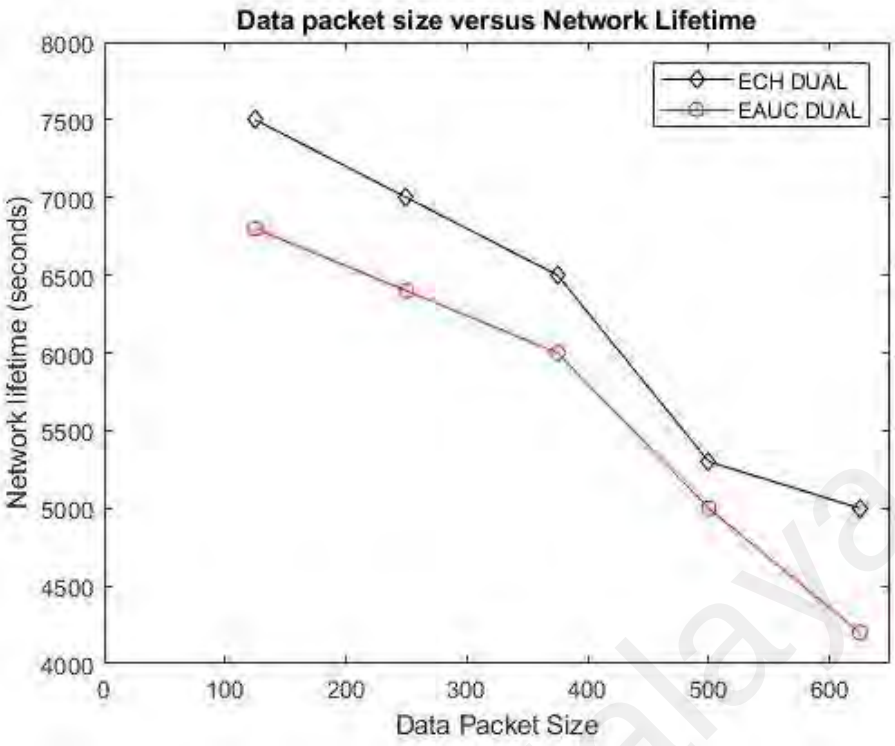

Figure 3.17: Network Lifetime versus Data Packet Size

Network lifetime is a function defined when first nodes run out of energy and refrains from forwarding. In figure 3.17 a decrease in network lifetime is observed with increasing packet size. ECH-DUAL has longer connectivity ensuring higher network lifetime than the EAUC-DUAL. Table 3.14 shows the comparison of data packet size and network lifetime of ECH-DUAL and EAUC-DUAL.

Table 3.14: Comparison of Data Packet Size and Network Lifetime

Data Packet Size	Network lifetime -ECH DUAL (seconds)	Network lifetime -EAUC DUAL (seconds)
125 B	7521	6843
250 B	7005	6407
375 B	6456	5924
500 B	5321	4939
625 B	4881	4237

3.7 Cox Regression Analysis

Cox regression denotes the occurrence of an event at a time. The survival rate prior and posterior to an event with respect to time is denoted as hazard ratio. This research work figures out the numerical values of hazard ratio for the proposed protocol. The impact of the results obtained and its numerical deviation from its conventional protocol is significant. SPSS has been used for numerical interpretation and there are prior works analyzing the results obtained and also depicts the relationship between signal strength and energy consumption (Xu et al., 2010). Impact of data generation with the metrics of delivery has been analyzed linearly with regression using the results obtained. Similarly traffic volume analysis is performed for clustering using SPSS tool for the raw data provided to form appropriate clusters (Kanget al., 2019).

Data analysis is performed on the results obtained from ECH-DUAL and EAUC- DUAL using Cox regression. The parameters used for analysis of the two algorithms are alive nodes and average residual energy. The data analysis using Cox regression is performed in order to validate the obtained simulation results of ECH-DUAL and EAUC-DUAL. Statistical analysis is also performed to prove the alternate hypothesis. Table 3.15 depicts the overall result of alive nodes.

Table 3.15: Overall results of Alive Nodes - MATLAB

Rounds	ECH – DUAL_A (Proposed)	ECH – DUAL (Status)	EAUC– DUAL_A	EAUC– DUAL (Status)
250	100	0	100	0
500	100	0	100	0
1000	100	0	28	1
1120	99	1	6	1
1500	70	1	4	1
2000	12	1	3	1
2500	5	1	1	1
4250	0	1	0	1

Null Hypothesis: There is no impact of dead nodes and average energy consumption with increase in rounds.

Alternate Hypothesis: There is an impact on number of dead nodes and average energy consumption with increase in rounds. The round indicates the time the simulation is being run. The hazard ratio is the death of node denoted by Exp (B) in table 3.20 and table 3.27.

Dependant variable: Rounds

Covariates: Alive nodes and average residual energy.

In table 3.8 the status 1 indicates the node is dead, the status 0 indicates the alive nodes. A single node death can partition the sensing field in a network.

3.7.1 Cox Regression of ECH_DUAL_A

ECH_DUAL_A represents the alive nodes in the algorithm. Table 3.16 shows the case processing summary where the total number of rounds is 8. In this case 3 rounds are censored wherein death of node has not occurred. The corresponding percentage of event occurred or not occurred is also tabulated. "a" in the table denotes the parameter "rounds" and it is a dependant variable.

Table 3.16: Case processing summary of ECH_DUAL_A

	Case processing summary	**N_s**	**Percentage (%)**
Cases available in analysis	Event[a]	5	62.5
	Censored	0	0
	Total	5	-
	Cases with missing values	0	0
	Cases with negative time	0	0
Cases dropped	Censored cases before the earliest event in a stratum	3	37.5
		0	0
	Total	3	-
Overall Total		8	100

Table 3.17: Omnibus test for model coefficients of ECH_DUAL_A

-2 Log Likelihood
9.575

Table 3.18: Iteration history[b] of ECH_DUAL_A

-2 Log Likelihood[a]	Coefficient
	ECH_DUAL_A
3.513	.049
2.613	.079
1.791	.135

Table 3.17 shows the model coefficient for ECH_DUAL_A at the beginning of the block with no predictors. The value of initial log likelihood function denoted for stratification by -2 Log likelihood is 9.575. The iteration history in table 3.18 indicates the change beginning from the initial estimates to last three iterations. In this table "a" denotes the beginning block number 0 and "b" denotes the iteration values to check whether at least one coefficient tends to infinity after 3 iterations.

The chi square change shown in table 3.19 indicates the predictors denoted in terms of -2 Log Likelihood times the model at previous stage and current stage. "a" denotes the beginning block number 1. From table 3.19 the test is statistically significant with Sig value less than 0.05. Hence the test of null hypothesis can be rejected and alternate hypothesis is proved.

Table 3.19: Omni bus test for model coefficients[a] of ECH_DUAL_A

.

-2 Log Likelihood	Overall (score)			Change from Previous Step			Change from Previous Block		
	Chi-Square	df	Sig	Chi-Square	df	Sig	Chi-Square	df	Sig
1.791	5.853	1	.016	7.784	1	.005	7.784	1	.005

The confidence of estimates (survival function) is shown in table 3.20 where estimate of the hazard is denoted by Exp (B). The hazard ratio is 1.144. The "B" denotes the Regression coefficient, SE denotes standard Error, df denotes degree of freedom. CI denotes the Confidence interval. The Wald can be calculated by using $\left(\frac{B}{SE}\right)^2$. The mean of covariates is shown in table 3.21. The covariates indicate the mean of alive nodes. The figure 3.18 shows the survival function for ECH_DUAL_A signifying the life time table decrease with increase in rounds.

Table 3.20: Variables of ECH_DUAL_A

ECH_DUAL_A	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI Exp(B)	
							Lower	Upper
	.135	.151	.799	1	.371	1.144	.852	1.537

Table 3.21: Covariates mean of ECH_DUAL_A

	Mean
ECH_DUAL_A	37.200

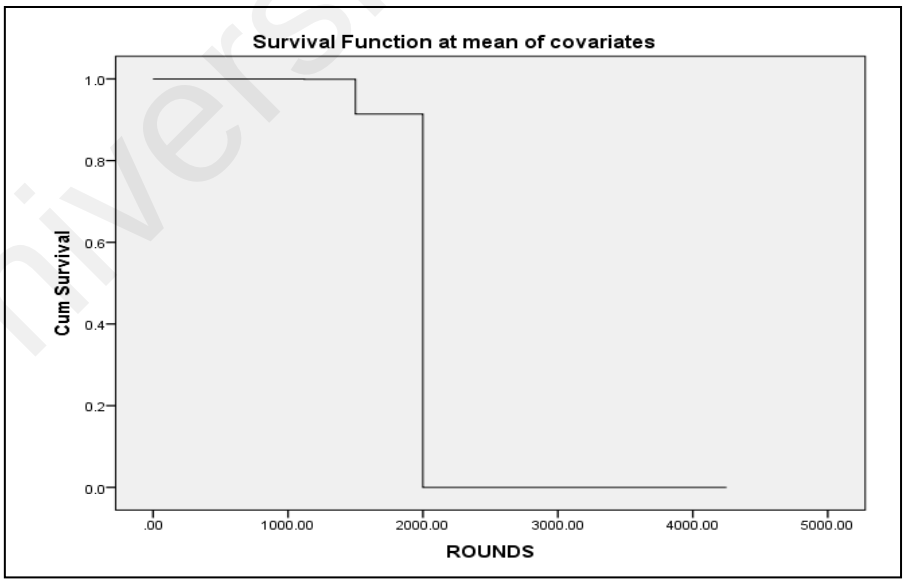

Figure 3.18: Survival Function of ECH_DUAL_A

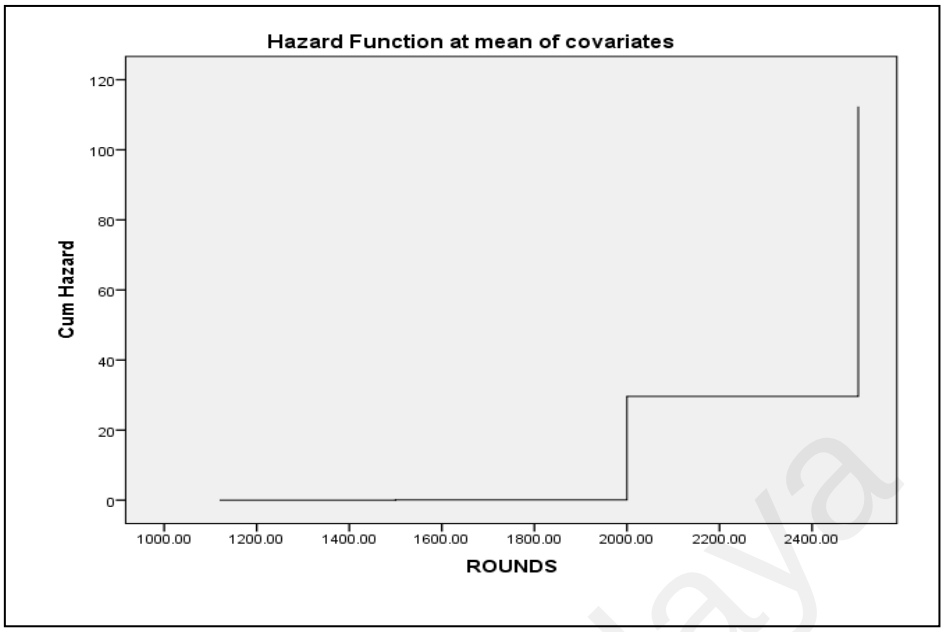

Figure 3.19: Hazard Function of ECH_DUAL_A

The figure 3.19 shows the Hazard function for **ECH_DUAL_A** signifying the risk factor increase with increase in rounds. The occurrence of risk factor is nearer to 1120.

3.7.2 Cox Regression of ECH_DUAL_AE

ECH_DUAL_AE represents the average residual energy in the algorithm. The table 3.22 indicates the simulation results obtained for both the proposed protocol and conventional protocol and the status of surviving nodes. Table 3.23 shows the case processing summary where the total number of rounds is 8. In this case 3 rounds are censored wherein death of node has not occurred. The corresponding percentage of event occurred or not occurred is also tabulated. “a” in the table denotes rounds as a dependent variable.

Table 3.22: Overall results of Average Residual Energy - MATLAB

Rounds	ECH _DUAL_AE (Proposed)	ECH_DUAL (Status)	EAUC– DUAL_AE	EAUC– DUAL (Status)
250	0.9	0	0.37	0
500	0.7	0	0.22	0
1000	0.4	0	0.02	1
1120	0.37	1	0.01	1
1500	0.10	1	0	1
2000	0.04	1	0	1
2500	0.01	1	0	1
4250	0	1	0	1

Table 3.23: Case processing summary of ECH_DUAL_AE

	Case processing summary	N_s	Percentage %
Cases available in analysis	Event[a]	5	62.5
	Censored	0	0
	Total	5	-
	Cases with missing values	0	0
	Cases with negative time	0	0
Cases dropped	Censored cases before the earliest event in a stratum	3	37.5
		0	0
	Total	3	-
Overall Total		8	100

Table 3.24 shows the model coefficient for ECH_DUAL_AE at the beginning of block with no predictors. The value of initial Log Likelihood function denoted by -2 Log likelihood is 9.575.

Table 3.24: Omnibus test for model coefficients of ECH_DUAL_AE

-2 Log Likelihood
9.575

Table 3.25: Iteration history[b] of ECH_DUAL_AE

-2 Log Likelihood[a]	Coefficient
	ECH_DUAL_AE
3.854	17.208
2.374	34.749
1.367	61.283

Table 3.26: Omni bus test model coefficients[a] of ECH_DUAL_AE

-2 Log Likelihood	Overall (score)			Change from Previous Step			Change from Previous Block		
	Chi-Square	df	Sig	Chi-Square	df	Sig	Chi-Square	df	Sig
1.367	6.140	1	.013	8.208	1	.004	8.208	1	.004

The iteration history in table 3.25 indicates the change beginning from the initial estimates to last three iterations. In this table "a" denotes the beginning block number 0 and "b" denotes the iteration values to check whether at least one coefficient tends to infinity after 3 iterations. The chi square change shown in table 3.26 indicates the predictors denoted in terms of -2 Log Likelihood times the model at previous stage and current stage. "a" denotes the beginning block number 1. From table 3.26 the test is statistically significant with Sig value less than 0.05. Hence the test of null hypothesis can be rejected and alternate hypothesis is proved. The "B" denotes the Regression coefficient, SE denotes standard Error, df denotes degree of freedom. The confidence of estimates (survival function) is shown in table 3.27 where in estimate of the hazard is denoted by Exp (B). The hazard ratio is 4.118E+26.

Table 3.27: Variables of ECH_DUAL_AE

ECH_DUAL_AE	B	SE	Wald	Df	Sig.	Exp(B)	95.0% CI Exp(B)	
							Lower	Upper
	61.283	54.088	1.284	1	.257	4.118E+26	.000	4.511E+072

The mean of covariates is shown in table 3.28. The covariate in this work indicates the average energy consumption of nodes. The figure 3.20 shows the survival function for ECH_DUAL_AE signifying the life time table decrease with increase in rounds. The figure 3.21 shows the Hazard function of ECH_DUAL_AE signifying the risk factor increases with increasing rounds. The occurrence of risk factor is nearer to 1120.

Table 3.28: Covariates mean of ECH_DUAL_AE

	Mean
ECH_DUAL_AE	37.200

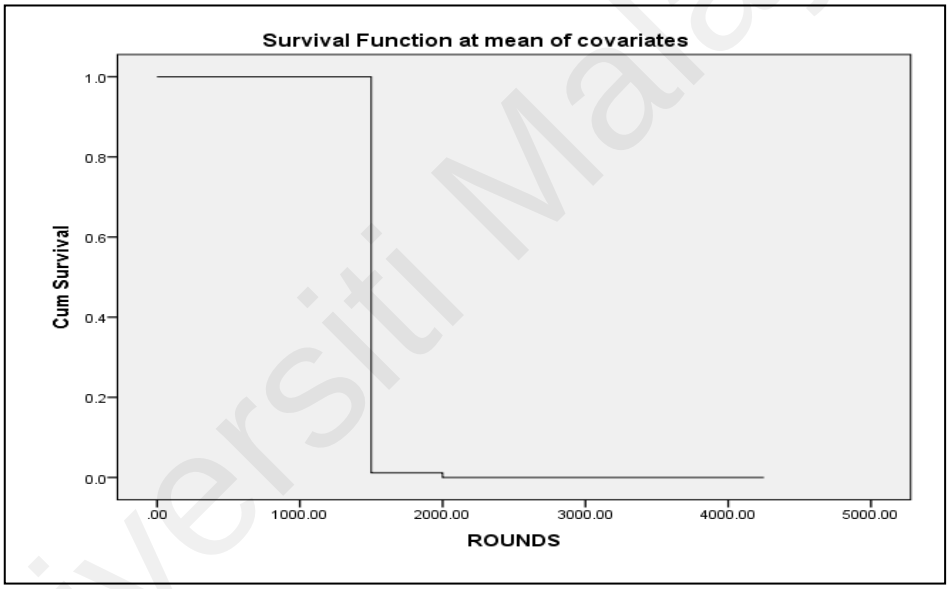

Figure 3.20: Survival Function of ECH_DUAL_AE

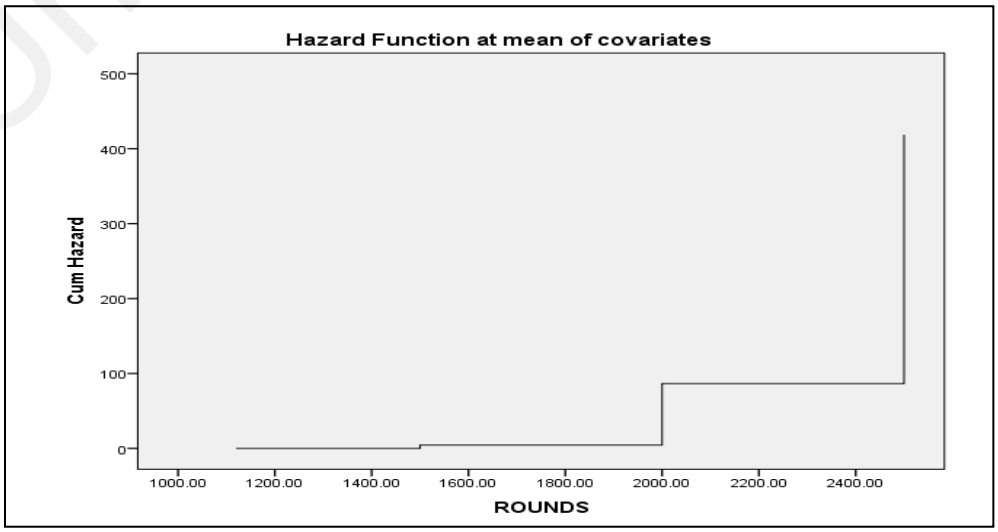

Figure 3.21: Hazard Function of ECH_DUAL_AE

3.7.3 Cox Regression of EAUC_DUAL_A

EAUC_DUAL_A represents the alive nodes in the algorithm. The table 3.15 shows the summary of alive and dead nodes of EAUC_DUAL_A. Using this the case processing summary for the algorithm is developed in table 3.29. The total number of rounds in table 3.29 is 8. The 2 rounds are censored wherein death of node has not occurred. The corresponding percentage of event occurred or not occurred are also shown. "a" represents the rounds as a dependent variable.

Table 3.30 shows the model coefficient of ECH_DUAL_A at the beginning of the block without any predictors. The value of initial Log Likelihood function denoted by -2 Log likelihood is 13.159. The iteration history in table 3.31 indicates the change beginning from the initial estimates to last three iterations. In this table "a" denotes the beginning block number 0 and "b" denotes the iteration values to check whether at least one coefficient tends to infinity after 3 iterations.

Table 3.29: Case processing summary of EAUC_DUAL_A

	Case processing summary	**N_s**	**Percentage %**
Cases available in analysis	Event[a]	6	75
	Censored	0	0
	Total	6	-
	Cases with missing values	0	0
	Cases with negative time	0	0
Cases dropped	Censored cases before the earliest event in a stratum	2	25
		0	0
	Total	2	-
Overall Total		8	100

Table 3.30: Omnibus test for model coefficients of EAUC_DUAL_A

-2 Log Likelihood
13.159

Table 3.31: Iteration history[b] of EAUC_DUAL_A

-2 Log Likelihood[a]	Coefficient
	EAUC_DUAL_A
6.137	.281
2.935	.750
1.336	1.319

The chi square change shown in table 3.32 indicates the predictors denoted in terms of -2 Log Likelihood times the model at previous stage and current stage. "a" denotes the beginning block number 1. The confidence of estimates (survival function) is shown in table 3.33. The estimate of the hazard is denoted by Exp (B). The hazard ratio is 3.739. From table 3.32 the test is significant as the Sig value is less than 0.05. Hence the test of null hypothesis is rejected. The mean of covariates is shown in table 3.34. The covariate in this work indicates the alive nodes.

Table 3.32: Omni bus test for model coefficients[a] of EAUC_DUAL_A

-2 Log Likelihood	Overall (score)			Change from Previous Step			Change from Previous Block		
	Chi-Square	df	Sig	Chi-Square	df	Sig	Chi-Square	df	Sig
1.336	7.978	1	.005	11.822	1	.001	11.822	1	.001

Table 3.33: Variables of EAUC_DUAL_A

EAUC_DUAL_A	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI Exp(B)	
							Lower	Upper
	1.319	.873	2.281	1	.131	3.739	.675	20.701

Table 3.34: Covariates mean of EAUC_DUAL_A

	Mean
EAUC_DUAL_A	7.000

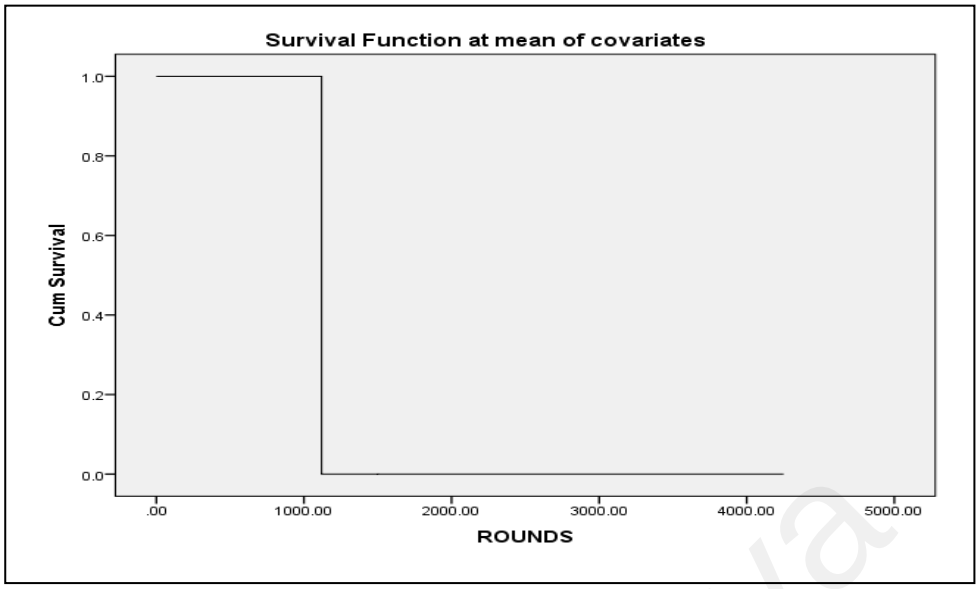

Figure 3.22: Survival Function of EAUC_DUAL_A

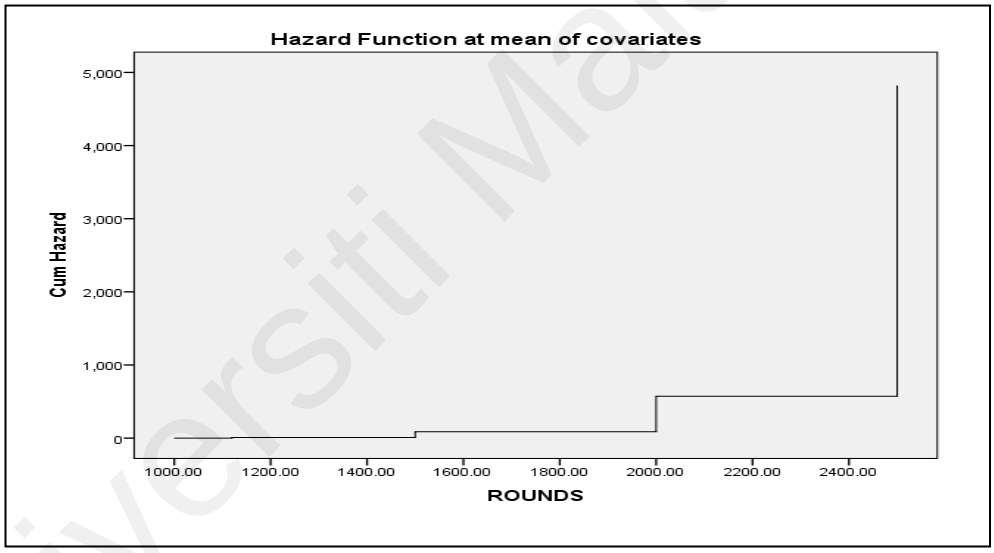

Figure 3.23: Hazard Function of EAUC_DUAL_A

The figure 3.22 shows the survival function for EAUC_DUAL_A signifying the network life time decrease with increase in rounds. The figure 3.23 shows the hazard function of EAUC_DUAL_A signifying the risk factor increase with increase in rounds. The occurrence of risk factor is nearer to 1000.

3.7.4 Cox Regression of EAUC_DUAL_AE

Table 3.35: Case processing summary of EAUC_DUAL_AE

	Case processing summary	N_s	Percentage %
Cases available in analysis	Event[a]	6	75
	Censored	0	0
	Total	6	-
	Cases with missing values	0	0
	Cases with negative time	0	0
Cases dropped	Censored cases before the earliest event in a stratum	2	25
		0	0
	Total	2	-
Overall Total		8	100

Table 3.36: Omnibus test for model coefficients of EAUC_DUAL_AE

-2 Log Likelihood
13.159

EAUC_DUAL_AE represents the average residual energy in the algorithm. The table 3.22 specifies the simulation results obtained for both the proposed protocol and conventional protocol and the status of surviving nodes. Table 3.35 shows the case processing summary in where the total number of rounds is 8. The 2 rounds are censored wherein the death of node has not occurred. The corresponding percentage of event occurred or not occurred are also shown. Table 3.36 shows the model coefficient of EAUC_DUAL_AE at the beginning of the block without any predictors. The value of initial Log Likelihood function denoted by -2 Log likelihood is 13.159.

The iteration history in table 3.37 indicates the change beginning from the initial estimates to last three iterations. In this table “a” denotes the beginning block number 0 and “b” denotes the iteration values to check whether at least one coefficient tends to infinity after 3 iterations

Table 3.37: Iteration history[b] of EAUC_DUAL_AE

-2 Log Likelihood[a]	Coefficient
	EAUC_DUAL_AE
6.794	309.417
6.512	415.092
6.412	517.583

Table 3.38: Omni bus test for model coefficients[a] of EAUC_DUAL_AE

-2 Log Likelihood	Overall (score)			Change from Previous Step			Change from Previous Block		
	Chi-Square	df	Sig	Chi-Square	df	Sig	Chi-Square	df	Sig
6.412	7.117	1	.008	6.746	1	.009	6.746	1	.009

Table 3.39: Variables of EAUC_DUAL_AE

EAUC_DUAL_AE	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI Exp(B)	
							Lower	Upper
	517.583	600.731	.742	1	.389	6.076E+224	.000	-

Table 3.40: Covariates mean of EAUC_DUAL_AE

	Mean
EAUC_DUAL_A	.005

The chi square change shown in table 3.38 indicates the predictors denoted in terms of -2 Log Likelihood times the model at previous stage and current stage. “a” denotes the beginning block number 1. The confidence of estimates (survival function) is shown in table 3.39. The estimate of the hazard is denoted by Exp (B). The hazard ratio is 6.076E+224. From table 3.39 the test is statistically significant as the Sig value is less than 0.05. Hence the test of null hypothesis can be rejected. The mean of covariates is shown in table 3.40. The covariate in this work indicates the average energy consumption of nodes. The figure 3.24 shows the survival function of EAUC_DUAL_AE signifying the network life time decrease with increase in rounds.

The figure 3.25 shows the Hazard function of EAUC_DUAL_AE signifying the risk factor increases with increase in rounds. The occurrence of risk factor is nearer to 1000. From the cox regression analysis, it is clearly evident where the null hypothesis is rejected and alternate hypothesis is proved in both the cases (parameters) of ECH_DUAL. Comparing the two protocols the ECH_DUAL proves better than the EAUC_DUAL in terms of Hazard ratio denoted by Exp (B).

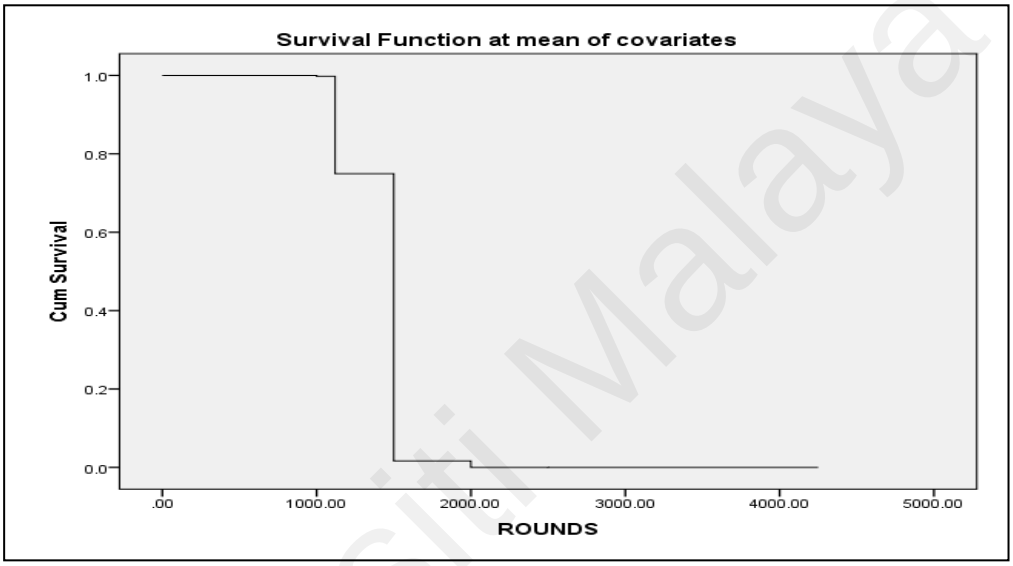

Figure 3.24: Survival Function of EAUC_DUAL_AE

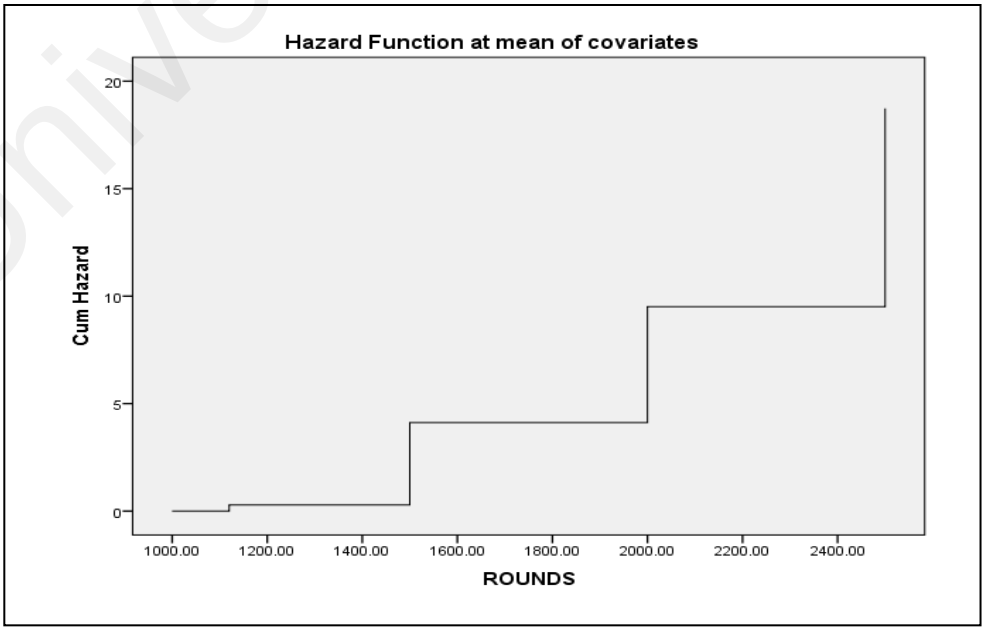

Figure 3.25: Hazard Function of EAUC_DUAL_AE

3.8 Conclusion

Energy Aware Unequal Clustering with Dual Sink, EAUC-DUAL suggests dual sink deployed in the network to gather data and prolong the lifetime of sensor nodes. The protocol minimized the delay and hotspot problems by integrating the advantages of unequal clustering and the data processing features of static and mobile sinks. EAUC-DUAL achieved load balancing among clusters with the formation of smaller-sized clusters near to the static sink while larger sized clusters farther away from static sink. Load balancing was also achieved with fixed and dynamic paths. By means of intra-cluster and inter-cluster routing approaches aggregated data were transmitted to the appropriate sink. EAUC-DUAL enhanced network lifetime and reduced energy consumption. The proposed protocol was found to perform better than the LEACH protocol with Static sink and LEACH protocol with DUAL sink.

A larger sensing filed with uneven clustering and a new cluster head selection method is proposed in ECH-DUAL to improve the network lifetime of WSN. An unbiased opportunity for cluster members to become cluster heads has been proposed in ECH-DUAL. The ECH-DUAL algorithm proposed tentative cluster head selection and final cluster head selection procedures to reduce the energy consumption. The rotation of tentative cluster heads and selection of final cluster heads yields a superior performance. Tentative cluster head is chosen based on energy-based timer and trust value. From the tentative CH final cluster head is selected using node degree, competition range and residual energy. The uneven clustering method with dual sink is employed to balance the energy consumed by cluster heads and improve node lifetime in wireless sensor networks. The sensor nodes communicate within the sensing field to transfer data to the static or mobile sink thereby providing increased network lifetime. The transmission from the cluster head to the sink through static or mobile sink based on the distance reduces interference. The simulation results of ECH-DUAL have been

compared with EAUC-DUAL in terms of the number of alive nodes, energy efficiency and network lifetime. The scalability of EAUC-DUAL and ECH-DUAL was tested for varying network sizes. Using Cox regression the obtained results are analyzed so as to interpret the energy consumption and the number of alive nodes. The results of the data analysis prove the alternate hypothesis is valid and there is a strong relation between energy consumption and alive nodes. In addition the hazard ratio is better in the case of ECH-DUAL compared to EAUC-DUAL.

CHAPTER 4: INACS - INTELLIGENT NEIGHBOUR AIDED COMPRESSIVE SENSING

4.1 Introduction

Compressive sensing technique samples the signal at a rate much below the Nyquist sampling rate by trusting on the sparsity of the signal. Signals can have sparse or compressible representation in original domain or in a transformed domain. Also the reconstruction algorithms of compressive sensing can devotedly reconstruct the original signal back from lesser compressive measurements. Thus physical data acquisition can be achieved below the sampling rate by incorporating sparsity in sensed physical data. Compressive sensing reduces the sensing rate enormously in dense wireless sensor nodes deployed within a sensing field. This is done by means of exploiting the sparsity of the signal and it provides fewer measurements during data transfer and also proper recovery of the original signal is achieved using the same fewer measurements. Once compressive sensing is performed the limitation occurs during data forwarding when all the spatial co-located sensors presume to sense the occurrence of an event and simultaneously transmit leading to collision. This process burdens the network resources.

The proposed Intelligent Neighbour Aided Compressive Sensing Algorithm (INACS) decides which sensor has to transmit the data to the immediate node within the sensing field thereby reducing the collisions and the burden of network resources. Data forwarding takes place in INACS using Pearson's correlation coefficient. This coefficient is formed based on spatial and temporal coordinates (intra-signal propagation time) of neighbor nodes. Thus when there is more than one neighbour node during forwarding then the sensor node simply sends readings within the sensing period to a uniquely selected neighbor based on the highest correlation.

In this algorithm the number of transmissions does not monotonically increase when the time period of sensing increases. INACS has been compared with compressive wireless sensing (CWS) technique. Time series analysis is also performed to validate the simultaneous association between the number of transmissions and the time period.

4.2 Previous Work on Compressive Sensing

In traditional compression techniques signals are obtained at the full Nyquist rate and then compressed for suitable representation. On the contrary compressive sensing with sub-Nyquist criteria explains, a small number of linear measurements with non-zero coefficients can provide signal recovery by satisfying the conditions of incoherence and the Restricted Isometry Property (RIP) (Chen et al., 2012). In the process of event monitoring to reduce the communication cost the vector representation corresponds to the magnitude of the event and the position coordinates (Tropp et al., 2006). Hence two algorithms are discussed: the first is the partial consensus algorithm obtaining data corresponding to its event; the second is the Jacobi-based approach describing the event monitoring with regard to its own position (Tropp et al., 2006). Thus to reduce the communication cost in a wireless environment with spatial and temporal characteristics sparsity in data along with recovery is a significant metric. Sparsity-driven decomposition has been classified into two types namely: "cluster coherence" and "cluster sparsity" (Donoho & Kutyniok, 2013). Cluster coherence does not depend on the arrangement of non-zero entities whereas there is a focus on the RIP and coherence. The cluster sparsity arrangement of non-zero entities also plays a role in recovery process.

4.2.1 Network Cost Incurred in Compressive Sensing

Reducing observational cost and transmission cost is a pivotal task. The large-scale sensor deployment and transfer of data have been achieved by dividing the entire monitoring area into smaller regions of appropriate transmission ranges (Fan et al., 2009). The capacity of transmission associated with compressive sensing in wireless sensor networks is discussed in lower and upper bounds (Zhang et al., 2009). Lower capacity and upper bound are denoted by the following formulas by (Zhang et al., 2009).

$$Lower\ Transmission\ rate\ per\ unit\ radius = \Theta\left(\frac{W_c}{M_c}\right)$$

$$Upper\ Transmission\ rate\ per\ unit\ radius = \Theta\left(\frac{\sqrt{n}W_c}{M_c\sqrt{\log n}}\right)$$

Where "M_c" denotes compressed signal "W_c" channel bandwidth. and "n" number of nodes (Zhang et al., 2009). Compressive sensing reduces the sampling rate by 25% if a full probabilistic Bayesian framework and a sparse event detection process are used in wireless sensor networks (Meng et al., 2009). The two criteria considered for sparse events are the ratio of active sensors to total sensors deployed and the nature of the event occurring due to the simultaneously superimposition of the signal (Meng et al., 2009).

Localizing a sparse target within a sensing field and incorporating compressive sensing provide high reconstruction accuracy. Sparse targets located at different positions result in varying energy decay factors; hence the measurement matrix can satisfy the RIP. The novel greedy matching pursuit ensures sparse recovery with target counting (Zhang et al., 2011). Sparsity is obtained during the process of sensing by turning off some sensors during the iteration. This is done by determining the signal strength between neighbor sensors using a consensus-based algorithm (Ling & Tian, 2010). Each active sensor has its own measurement and it makes the recovery process difficult while assigning the desired communication range based on nodal density. Iterative hard thresholding (IHT) in compressing sensing is used to guarantee recovery with minimal errors and a fixed number of iterations (Blumensath & Davies, 2009). The IHT process is established with a sampling operator and its transpose function. The two versions of distributed computing with IHT are discussed: one for a static network and the other for a time-varying dynamic network. Static networks use the local computation represented by projection vector followed by global computation.

In dynamic networks the inexact computation of values is obtained and it is refined using a consensus-based algorithm with a weight function, thereby reducing computational time and bandwidth (Patterson et al., 2014). The sparse binary matrix considers the spatial coordination between sensors in data gathering (Lv et al., 2019). A numerical sparsity-based binary matrix achieves better recovery performance than conventional transforms such as discrete cosine transform (DCT) and discrete Fourier transform (DFT). It is also suitable for certain applications assuming that, time complexity involved in achieving numerical sparsity is high.

4.2.2 Reduction of Transmissions in Compressive Sensing

Minimum Energy Compressed Data Aggregation (MECDA) is for small and large-scale wireless sensor networks. In MECDA after assigning aggregation, routing provides an optimal solution in terms of energy efficiency (Xiang et al., 2011). The topology is configured to the aggregation set and the forwarding set. The non-uniform compressive sensing (NCS) method exploits heterogeneity in the transmission energy of sensors and non-uniform sampling. NCS delivers signal accuracy and a reduction in energy consumption by reducing the number of samples more effectively than conventional compressive sensing (Shen et al., 2013). Reducing the communication cost in wireless sensor networks has been achieved with spatially available sensory data (Wu et al., 2014). In other words the sparest solution of the measurement matrix should complement the requirement of the sensors resulting in considerable energy savings in wireless senor networks. The reduction in energy consumption is central to Adaptive Non-uniform Compressive Sampling (ANCS). It focuses on the corresponding coefficient of sensing energy. It achieves minimal error when sparse representation is available on a canonical basis (Zaeemzadeh et al., 2017). Reducing the number of transmissions has also been achieved with a proper measurement matrix using a Markov chain and compression probability (Huang & Soong, 2019).

4.2.3 Compressive Sensing and Joint Routing Capabilities

Compressive Data Gathering (CDG) together with compressive sensing has been proposed. CDG is centered on the partial transfer of data using the weighted sum of sensor readings. It as well considers the topology between the parent and the child nodes for forwarding while imposing a burden on specific nodes leading to hotspots (Luo et al., 2009). Distributed data gathering and compressive sensing have been discussed in relation to Principal Component Analysis (PCA) (Masiero et al., 2009). The significant role of PCA is the transformation matrix provided with an appropriate sparsity level. *Sample mean* and *covariance matrix* are the two terms considered for reconstruction quality in PCA. A Spatial Temporal Compressive Data Gathering scheme (STCDG) has been proposed to achieve a reduction in traffic levels. The recovery accuracy is higher due to the methodology of using stability in a short time along with a low-rank matrix. This approach is better than one-dimensional compression (Cheng et al., 2013). Incorporating fuzzy-based routing with CDG, localization and a geographic routing technique has been discussed (Ghaderi, et al., 2019).

Fuzzy logic is used in the selection of cluster heads and also involved in the decisions of the routing table to determine the hop count. Therefore the metrics for reduction in energy consumption and improved lifetime have been achieved by increasing the sensing field more effectively than conventional geographical routing. Compressive sensing with single sink and multiple sink architecture has been discussed. The capacity depends on two main metrics: data rate and random projections. A pipeline scheduling scheme is considered in the case of a single sink. In scenarios with multiple sink a single random projection is chosen out of a number of sensors within a session and thus capacity for gathering data is determined (Zheng et al., 2013). Hybrid Compressive Sensing (HCS) uses topological information to reduce the intra- and inter-

sensor communication to the sink. The cluster members transfer the data to the cluster head without compressive sensing, whereas the cluster head incorporates compressive sensing and transfers the data. The projection is calculated in HCS by the cluster head using a number of sensor nodes with the desired sparsity level as explained in (Xie & Jia, 2013). A sparse hybrid compressive sensing method has been proposed. In this approach the sink communicates the seed vector to the cluster head. The cluster head then generates a sparse matrix to communicate the measurement value to the sink. This method achieves the desired sparsity level and energy consumption. (Zhang et al., 2018).

4.2.4 Spatial and Temporal Compressive Sensing

The three-step process for compressive sensing with a discussion on the measurement matrix is presented in (Arjoune et al., 2018). The foremost step is sparse representation followed by linear encoding of collected measurements. The last step is the sparse recovery of signals by non-linear encoding. A data loss pattern in wireless sensor networks has been modeled with an environmental matrix incorporating missing values (Kong et al., 2014). This permits for higher reconstruction accuracy to be achieved with temporal stability when similarities in sensor readings are prevalent (Kong et al., 2014). A sensing matrix with a chaotic stream matrix using the features of zone matching and feedback control has also been discussed. A sensing matrix with a chaotic stream provides sampling efficiency with minimal complexity. Meanwhile a Topologically Conjugate Chaotic System (TCCS) achieves higher reconstruction accuracy in noisy environments (Gan et al., 2018). A Structurally Randomized Matrix (SRM) has been proposed where the signal is pre-randomized followed by fast transform and finally sub-sampling of measurements is performed (Do et al., 2012). Distributed source coding and compressive sensing is discussed involving two-sensor system. The distortion rate has been optimized with vector quantization by segregating

the noise component. It achieves a reduction in the mean square error during signal reconstruction (Wang et al., 2012). Neighbor-Aided Compressive Sensing (NACS) has been proposed with spatio-temporal correlation in wireless sensor networks (Quan et al., 2016). However NACS is a centralized approach relying on a single transform matrix.

The impact of MAC layer spatial correlation and distortion rate has been discussed (Vuran & Akyildiz, 2016). Two version of MAC protocols were being analyzed namely: event MAC and network MAC. In the case of the event MAC protocol the distance from source to event simultaneously increases the distortion rate, while the network MAC protocol prioritizes the packets during transmission depending on the occurrence of the event. Data Density Correlation Degree (DDCD) refers to the clustering of the sensor based on the amount of correlated data. Hence if high correlation is exhibited data can be grouped in the same cluster or grouped in different clusters (Yuan et al., 2013). The imputation of missing sensor readings has been performed with spatial-temporal processing (Li & Parker. 2014). Durbin Watson analysis is used to interpret whether or not time correlations were exhibited between sensors and Pearson's correlation revealed the spatial correlations between sensors.

Inter-node distance-based similarity in data measures has been discussed with regard to Euclidean distance, cosine similarity and Pearson's correlation. The results suggest, based on the application of sensors energy conservation can be achieved (Dhimal & Sharma, 2015). As the sensed signal is found to be higher in space and time it becomes enormous and has to be compressed. Temporal compression mostly relies on the network topology usually preferred for real time applications (Duarte et al., 2012). Link quality with transitional regions does not depend on distance; rather it is unpredictable and unreliable (Baccour et al., 2012). In a short time span the reception rate may be higher or lower due to constructive or destructive interference respectively. Weighted Spatial-Temporal Compressive Sensing (WST-CS) has been discussed and

found to ensure reconstruction quality by using weight vectors. The weight values lie between "0" and "1" where 0 indicates a small value of similarity and 1 indicates a high value of similarity. WST-CS ensures an increase in wireless sensor lifetime and accuracy (Mehrjoo & Khunjush, 2018). Coalition-based data gathering as well as a coalition coordinator is discussed in (Masoum et al., 2018). Certain sensors adjust the sampling rate during the local compressive sensing based on the coalition index as provided by the coalition coordinator. Reconstruction accuracy is achieved using two-step belief propagation-based recovery. Exhibiting sparsity in both temporal and spatial domains has been investigated (Wang et al., 2016). A data matrix sparse in both domains has higher reconstruction accuracy and improves the lifetime of the sensor network.

4.2.5 Transforms Deployed in Communication

Two categories of network data transform for wireless sensor networks have been stated (Shen et al., 2009). The first depends on the nature of data and is referred to as data-dependent while the other is termed as structure-dependent. In Data-dependent the network is adaptive and involves a learning rate while structure-dependent denotes the network is transformed into a non-adaptive one and does not involve a learning rate. "Invertible en-route in-network transform" has been discussed in relation to wireless sensor networks with unidirectional transmission (Shen & Ortega, 2010). In this process the source transmits the compressed data and the intermediate node decodes and adds its own data. However the initialization and scheduling of nodes make the process tedious. In the linear transform network a compression estimation matrix is formed based on the correlated signal flow between the source and relay (Goela & Gastpar, 2012). Network communication cost is reduced with approximate reconstruction. A cut in set bounds is obtained based on the factors of convex optimization and information theory.

4.2.6 Impact of Sampling on Compressive Sensing

In compressive sampling the appropriate classification of the signal and its subclass is mandatory to ensure reconstruction (Haupt et al., 2006). Sensing, Compression and Recovery through online Estimation (SCORE) has also been discussed whereby the raw signal transmitted at the source after random sampling is recovered at the destination at a data collection point. Depending on the estimated error, feedback is provided to align the random sampling rate (Quer et al., 2012). Distributed Adaptive Compressive Sensing (DASS) was proposed describing the design of an optimal sensing process using a temporal pattern. The process learns the signal statistics from past data. This is done by adjusting the pattern of sampling according to the signal statistics and finally confining the samples associated with these statistics. It achieves better efficiency than traditional compressive sensing in terms of spatial and temporal coordinates (Chen et al., 2012). Cost-Aware Compressive Sensing (CACS) interprets the sample cost and its associated recovery. Recovery guarantees are assured by using the Regularized Column Sum (RCS), specifying the lowest cost for recovery. Thus the cost of sampling is incorporated based on network resources (Xu et al., 2015).

A decentralized Bayesian Compressive Sensing (BCS) with a joint sparsity model is discussed along with intra and inter signal characteristics to reduce the burden at the fusion center (Chen & Wassell, 2015). The algorithm works by splitting the output into common and innovative components when the set of signals is given to a dictionary. The Frechet mean approach for calculating sparsity is performed using the signal weight. The non-zero position is a significant factor obtained with (PMP) Precognition Matching Pursuit (Chen et al., 2012). Meanwhile using a Sampling Rate Indicator (SRI) in the course of examining compressive sensing in wireless sensor networks was found to reduce the number of samples by 25% more than the conventional sampling method used in wireless sensor networks. This is achieved in

compressive sensing because the sampling frequency relies on characteristics of the signal rather than its bandwidth (Chen & Wassell, 2015). However compressive sensing is confined to the deployment of a wireless sensor network. Graph theory-based compressive sensing suggests, for reconstructing the signal vector for "n" connected nodes with a "k" sparse link the following equation can be used: O (k log (n)) (Xu et al., 2011). Reducing the sampling rate of individual sensors can be achieved based on spatial correlation and by using data prediction models at the sink. The lifetime of sensor networks can be increased by subsequently reducing the data transmission volume (Tayeh et al., 2019).

4.2.7 Network Model and Preliminaries

The network model deployed in this work uses a single sink with multi-hop communication. It is the responsibility of the sink to gather the information within the sensing period from the sensors. All the sensors have a transmission range identical with sensor nodes connected to one another. Observed data from a wireless sensor may be obtained at infrequent or frequent intervals and may require multiple dimensions depending on the sample rate (Duarte et al., 2012). But the enormous amount of data has to be considerably reduced due to the process of compression performed below the Nyquist rate. What is especially appealing is, spatial and temporal coordinates associated with varying sensors' data generation rate should sufficiently match with the session for transmission. Spatial coordinates indicate the position of the node, its transmission range and the intermediate forwarders. Temporal coordinates indicate the channel characteristics at an instant of forwarding taking into consideration the time interval needed to compress the observed sensory data. Designing a suitable structure in the current traffic scenario exhibits sparsity and recovery approximation is essential for compressive sensing in wireless sensors.

INACS considers unidirectional transmissions to avoid path mergers. Hence the work on INACS is compared with similar unidirectional transmission protocol called CWS (Bajwa et al., 2006). CWS considers source and channel communication in order to avoid path mergers in large-scale fading channels. The simulation of CWS has been performed by incorporating the same time period as INACS and as well using a large-scale fading channel.

4.3 INACS – Intelligent Neighbor Aided Compressive Sensing

In the previous work ECH- DUAL data aggregation is performed based on the Nyquist sampling theorem. This implies the sampling frequency should be twice as greater as the largest frequency at source for proper recovery of the signal. This process results in redundant data to be produced. Hence further algorithms on data gathering and data transfer will incorporate compressive sensing and it tends to reduce the data aggregation cost. INACS performs compressive sensing at the source and recovers the signal at the receiver using fewer measurements. The primary contribution of INACS is load balancing across the link during data forwarding. Proper routing coordinates are provided for each node to transmit data to the selected immediate node, based on correlation. This is achieved by an intelligent framework with a linear combination of two nodes (current forwarder and neighbor). The metrics taken are based on spatial and temporal coordinates (intra-signal propagation time) using Pearson's correlation coefficient.

At the source node the Rademacher matrix is formed based on Pearson's correlation with intertwined factors of spatial-temporal coordinates within the topology of routing. During the forwarding of compressed data when the source node has more than one neighbor node then it uses a criterion for selection of a neighbor node for routing. The criterion of selection is based on the neighbor node linearly associating with the present forwarder (source node) of highest correlation. Finally the usage of convex method improves the reconstruction accuracy of INACS at the receiver. The algorithm assumes for no abnormal reading affecting data fidelity. Secondly INACS considers the sparsity of the signal to be minimal with sensor nodes synchronized with the sink.

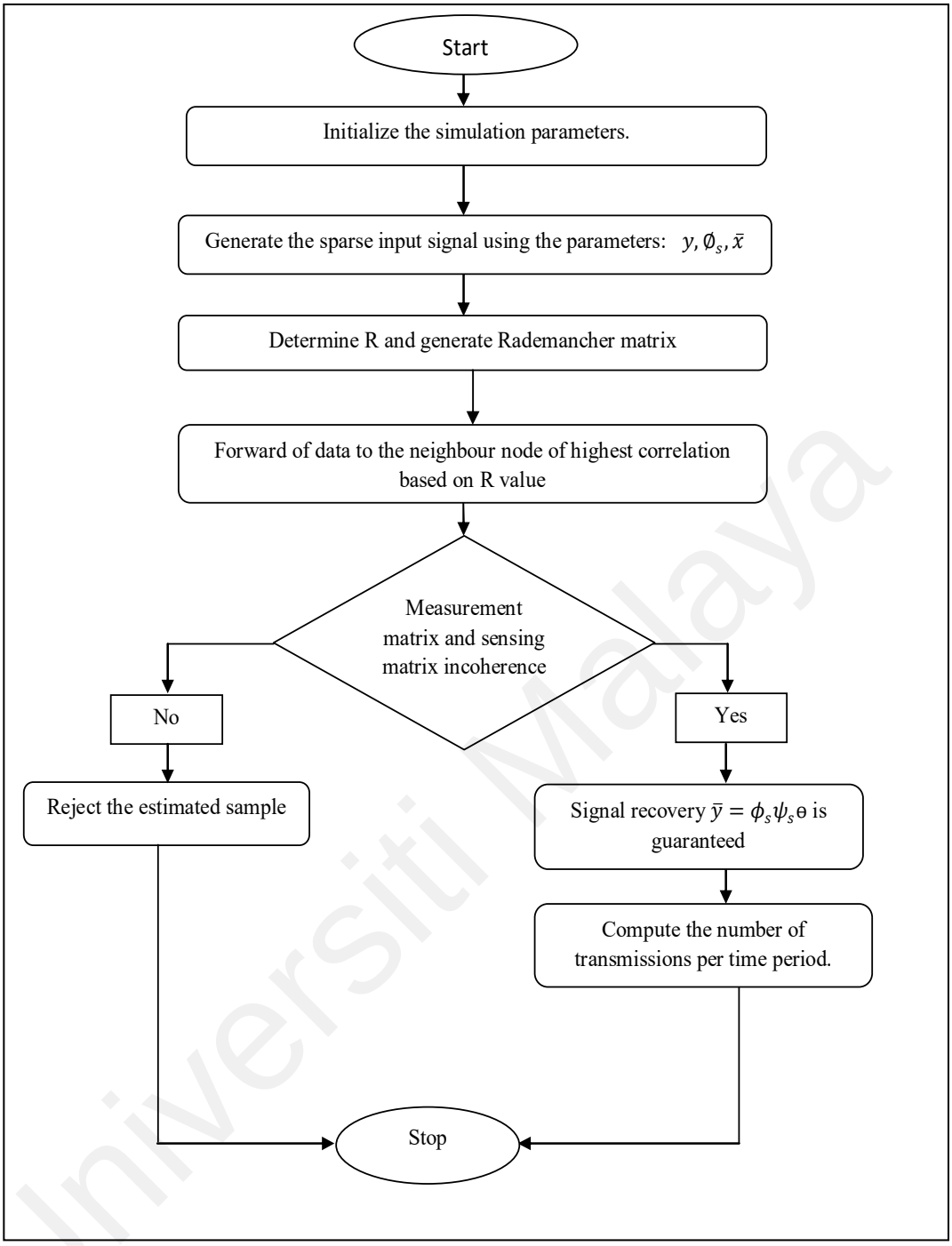

Figure 4.1: Flow chart of Intelligent Neighbor Aided Compressive sensing

Figure 4.1 shows the step by step process of initialization, signal generation, data forwarding, signal recovery and computation of Intelligent Neighbor Aided Compressive Sensing Algorithm (INACS). Thus the first step is to generate a sparse signal before the compressive sensing process. The second process after compressive sensing is data forwarding. This process generates the correlation coefficient "R" between the source node and each of its neighbor node. Using the correlation coefficient

a Rademacher matrix is formed at the source node. Thus the correlation information of itself (source node) and each of its neighbor node is made available at the source node. The source node can also be termed as "present forwarder". Now the source node decides to choose one of the neighbor nodes with highest correlation for forwarding the compressed data and rejects the transmission to other neighbor nodes. This eventually reduces the number of transmissions and hence overall the energy is saved in the network. The final process after data forwarding is signal recovery at sink. For the exact reconstruction at the receiver the sensing matrix and observation matrix should be incoherent. If this condition is not satisfied then the sample obtained at the destination will be rejected. Once the In-coherent condition is satisfied then the signal recovery is guaranteed. The number of transmissions per time period is computed.

4.3.1 INACS Algorithm for Gathering at Source

INACS employs an intelligent framework within a time period for correlating the spatial and temporal coordinates in the transmission of data to the sink. The "x" is an unknown signal with N dimension. It has to be compressed and transmitted.

$$\bar{y} = \phi_s \bar{x} \tag{4.1}$$

For compressing "x" equation 4.1 can be used. $\bar{x}$ denotes the sparse signal. The signal "x" denotes the spatial and temporal coordinates for the data, $\bar{y}$ denotes the observation matrix and ϕ_s denotes the sensing matrix.

$$\begin{vmatrix} y_1 \\ y_2 \\ y_M \end{vmatrix} = \begin{vmatrix} \phi_1 \\ \phi_2 \\ \phi_M \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \\ x_N \end{vmatrix} \tag{4.2}$$

M denotes the number of measurements. N denotes dimensions of the unknown signal. Signal correlation can be obtained for unknown signal using R

$$\boldsymbol{R} = \frac{\sum(\boldsymbol{A_i}-\bar{\boldsymbol{A}})(\boldsymbol{B_i}-\bar{\boldsymbol{B}})}{\sqrt{\sum(\boldsymbol{A_i}-\bar{\boldsymbol{A}})^2}\sqrt{(\boldsymbol{B_i}-\bar{\boldsymbol{B}})^2}} \tag{4.3}$$

Here in this work l_1 norm is used and enforces sparsity. Signal correlation can be obtained for unknown signal as in equation 4.3. The correlation measure (R) shows the association among and between spatial and temporal coordinates of nodes. Spatial coordinates are denoted by A_i, represented as the distance between the nodes (current forwarder and neighbor), and temporal coordinates (intra signal propagation time) are denoted by B_i. The value of $\bar{A}$ is the mean of spatial coordinates. The value of $\bar{B}$ denotes the mean of temporal coordinates.

The value of Equation 4.3 lies between -1 and +1. A positive value indicates higher correlation and a negative value indicates a lower correlation. A Rademacher matrix is formed by rounding off to +1 and -1 based on the value obtained in equation 4.3. The significance of the Rademacher matrix is, the coefficients are uniformly sampled between +1 and -1. Thus a sparse representation of the signal is obtained for forwarding with the highest association. Enabling a pairwise coefficient of spatial-temporal coordinates with routing results in higher reconstruction accuracy. Figure 4.2 depicts the scenario where sensors communicate to the sink via a multi-hop involving spatial and temporal routing.

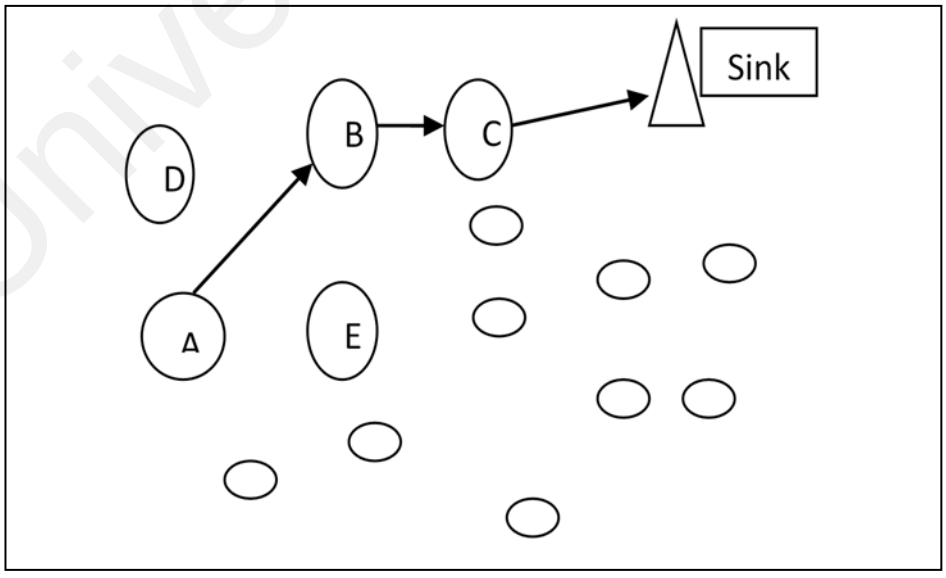

Figure 4.2: A scenario of sensor forwarding data to the sink

The source node is “A” and the eligible forwarders are B, D and E, as they are within the transmission range of A. The source node A prefers the B node due to a higher correlation coefficient. Simultaneously it suppresses the transmissions to other forwarders. Compressive sensing with structured random matrices provides higher accuracy with l_1 normalization (Blanchard et al., 2011). Thus the random signal “x” is made symmetrical to the values of “+1” and “-1” for the desired sparsity level thereby ensuring recovery.

4.3.2 Recovery Structure

The recovery structure for under-sampled rate and oversampled measurements is provided by the RIP (Blanchard et al., 2011). The orthogonal matrix as in the case of linear algebra states the real entries of rows and columns are orthogonal vectors and can be represented by an identity matrix (I) (Abo - Zahhad et al., 2015).

$$A^T A = AA^T = I \tag{4.4}$$

Enforcing sparsity can be done as in equation 4.5. Equation 4.5 can also be written in epigraph. Since, it comes under convex optimization technique.

$$min\|x\|_1 \tag{4.5}$$

The measurement matrix incoherence can be written by equation 4.6. ψ_s denotes the basis matrix.

$$A = \phi_s \psi_s \tag{4.6}$$

Thus the above equation is rewritten as in equation 4.7

$$\bar{y} = \phi_s \psi_s \theta \tag{4.7}$$

4.4 Results and Discussions

Simulation has been done with MATLAB software where in 50 sensor nodes are associated with transfer of data to single sink. The increase in dimension of wireless communication causes an increase in complexity**.** So implementation of simulation has been carried out till 10000 periods. Further statistical analysis has been made with time series analysis for associating the reduction in transmission cost achieved over the specific duration in section 4.5. Table 4.1 and table 4.2 depicts the numerical values of packets transmitted in CWS and INACS for varying transmission period or sensing period. The reduction in number of transmissions for each period of time is represented in percentage in table 4.2.

Table 4.1: Transmitted Packets versus Transmission Period of CWS

Protocol	Transmissions in 1000 periods	Transmissions in 5000 periods	Transmissions in 10000 periods
CWS	0.476×10^6	1.467×10^6	2.192×10^6

Table 4.2: Transmitted Packets versus Transmission Period of INACS

Protocol	Transmissions in 1000 periods	Transmissions in 5000 periods	Transmissions in 10000 periods
INACS	0.179×10^6	0.902×10^6	1.079×10^6
Percentage of reduction in transmissions	2.659	1.626	2.031

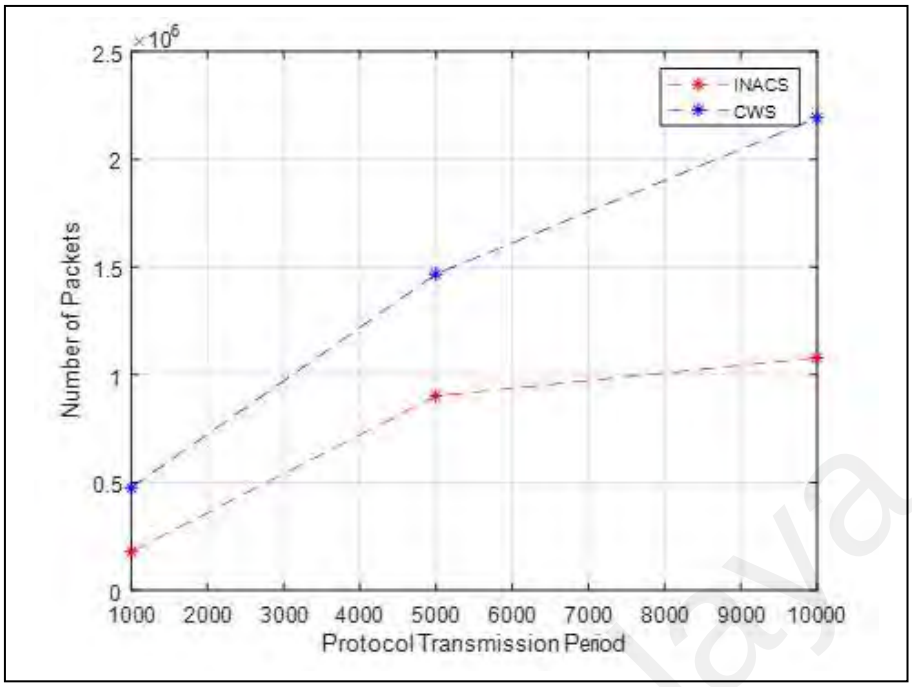

Figure 4.3: Number of Packets versus Protocol Transmission Period

Figure 4.3 shows the number of packets and the corresponding transmission periods. The number of transmissions for each period in INACS is reduced compared to CWS. Thus the energy consumption is reduced in INACS than in CWS enhancing the network lifetime. The simulation values of energy consumption are tabulated against scalability (number of nodes) as shown in table 4.3.

Table 4.3: Numerical values of Number of Nodes versus Energy Consumption

Number of Nodes	Energy consumption INACS	Energy consumption CWS
10	4.02 J	4.31 J
20	5.47 J	6.32 J
30	6.72 J	8.26 J
40	8.72 J	9.97 J
50	10.32 J	12.02 J

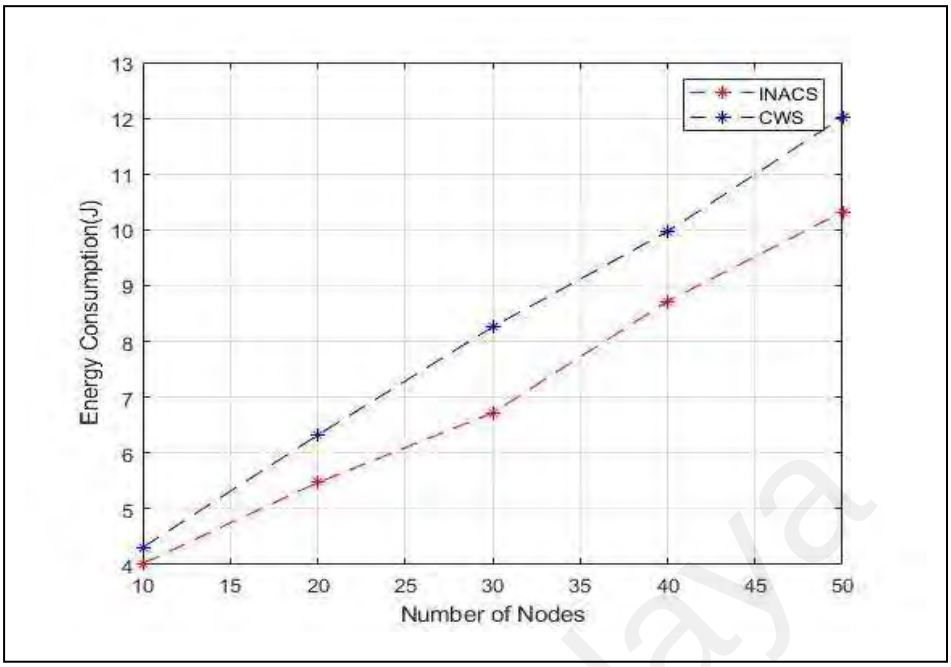

Figure 4.4: Number of Nodes versus Energy Consumption

Figure 4.4 shows the energy consumption for varying number of nodes. The energy consumption of INACS is considerably reduced than the conventional CWS. The figure 4.5 and table 4.4 shows the compression ratio and reconstruction error. The reconstruction accuracy of INACS is considerably higher than the conventional CWS. Number of packets generated at a source sensor at a particular time instance is denoted by S_i and the recovery of data at sink is denoted by S_d.

Reconstruction error has been calculated as in equation 4.8

$$Reconstruction\ Error = S_i - S_d \qquad (4.8)$$

Table 4.4: Compression Ratio versus Reconstruction Error

Compression Ratio	Reconstruction Error INACS	Reconstruction Error CWS
0.2	1	2
0.4	4	6
0.6	5	7
0.8	7	9
1.0	8	9

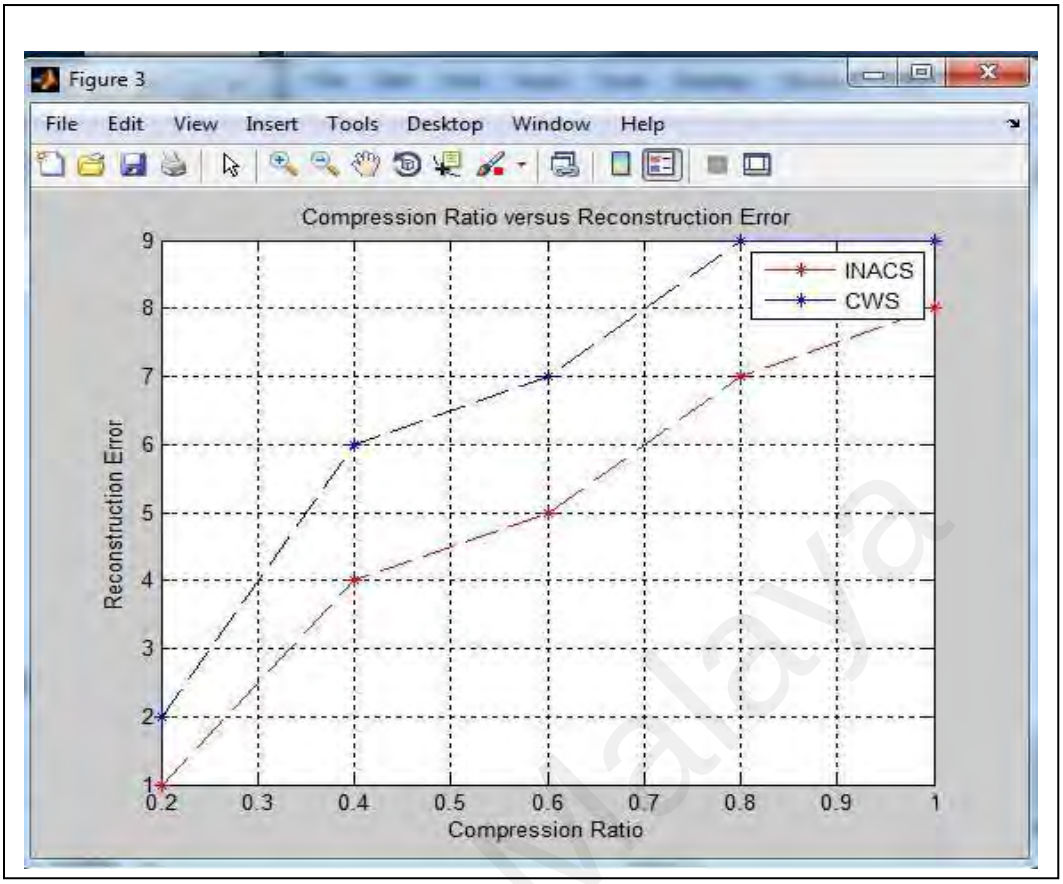

Figure 4.5: Compression Ratio versus Reconstruction Error

4.4.1 Scalability of INACS

Simulation has been done with MATLAB software where in 100 sensor nodes are associated with transfer of data to single sink for the algorithm INACS. The increase in dimension of wireless sensor network causes an increase in complexity**.** So implementation of simulation has been carried out till 10000 periods. The terrain size is also increased to 2000 m × 2000m. The transmission radius is varied from 20m to 100m. Thus the network size, the number of sensor nodes and the transmission radius is increased as in table 4.5 and the results are analyzed.

Table 4.5: Simulation Parameters with increased network size – INACS

Parameters	Values
Number of sensor nodes	100
Transmission radius	20 m to 100 m
Initial energy of sensor nodes	2 J
Terrain	2,000 m × 2,000m
Total duration of sensing periods	10,000 s

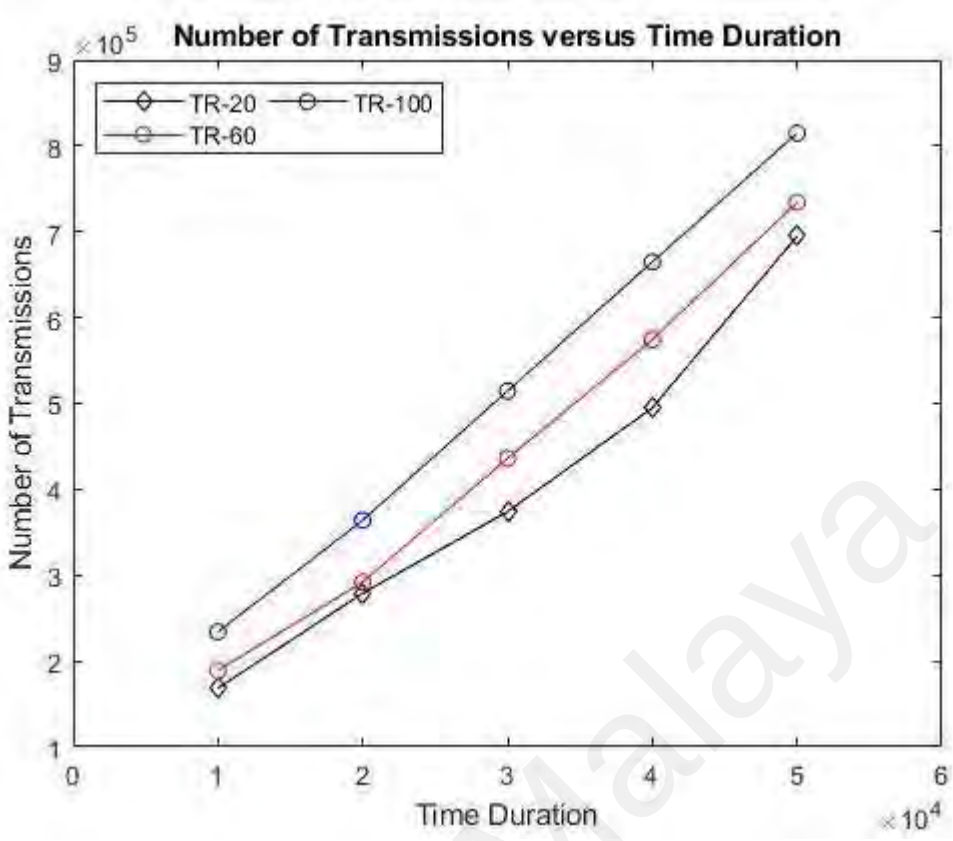

Figure 4.6: INACS - Number of Transmissions versus Time Duration.

Figure 4.6 shows the number of transmissions versus time period in seconds. Transmission radius is gradually increased for INACS. For smaller transmission radius the proximity of forwarding nodes is less and reduced transmission occurs. Larger transmission radius reduces the channel outage and increases the number of transmissions. Table 4.6 depicts the number of transmissions occurring for varying time period and varying transmission radius for INACS.

Table 4.6: Comparison on Number of Transmissions for varying Transmission Radius and Time Duration of INACS

Time duration in seconds	Number of Transmissions		
	(TR=20)	(TR=60)	(TR=100)
10000	168410	189100	234010
20000	279000	292000	364010
30000	374010	436000	514500
40000	494845	574300	664800
50000	694845	734300	814845

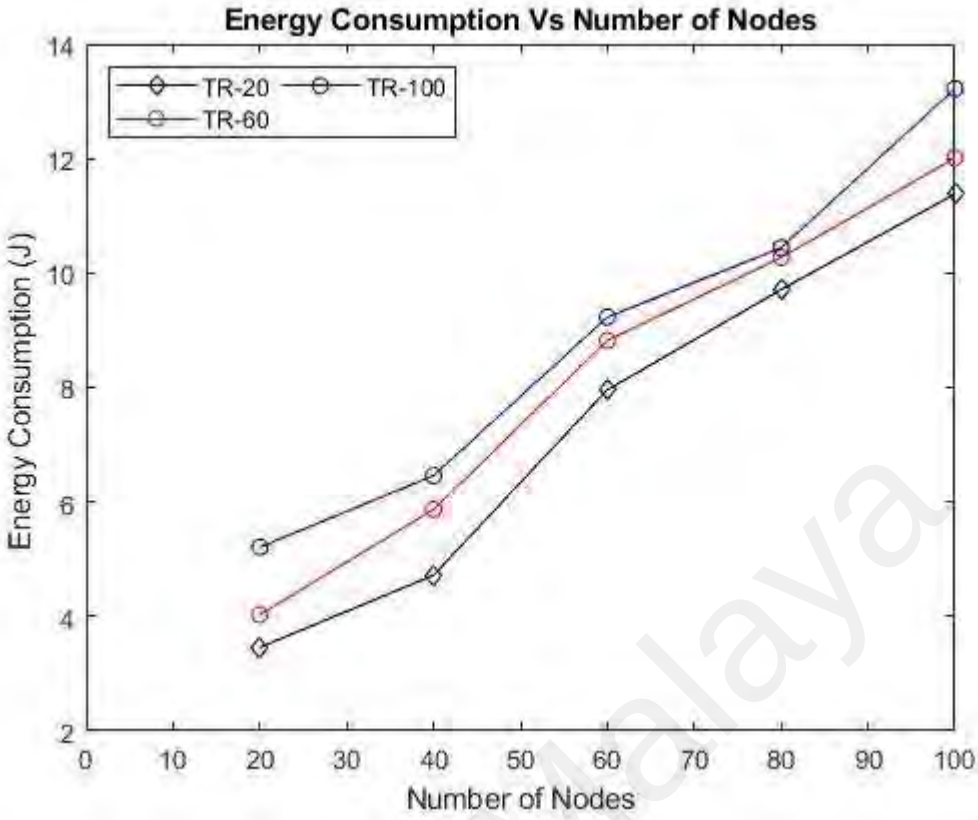

Figure 4.7: Energy Consumption versus Number of Nodes

Figure 4.7 shows the energy consumption versus number of nodes ensuring scalability. The transmission radius is gradually increased in INACS and more energy is consumed with increased number of nodes. Balancing energy efficiency is achieved using suitable transmission power in INACS. Table 4.7 depicts the energy consumed for varying transmission radius and varying number of nodes for INACS.

Table 4.7: Comparison on Energy Consumption for varying Transmission Radius Nodes of INACS

Number of Nodes	Energy Consumption		
	(TR=20)	(TR=60)	(TR=100)
20	3.46 J	4.03 J	5.21 J
40	6.72 J	5.87 J	6.46 J
60	7.96 J	8.82 J	9.23 J
80	9.71 J	10.27 J	10.44 J
100	11.38 J	12.01 J	13.22 J

4.5 Data Analysis of INACS and CWS

Time series forecasting is used for estimating the influence of position coordinates in localization using auto regressive moving average (Liu et al., 2015). However in this approach sink provides the previous target position where in the next target position is being calculated. The necessity of outlier detection for sensor local reading is analyzed to identify the model fits the data (Tulone & Madden, 2006). In several works obtaining the training data has been found to be tedious and thus achieving analytical performance for the obtained results is mandatory.

Monotonically decreasing energy consumption analysis has been done in time intervals using Auto regressive Integrated Moving average (ARIMA) model in SPSS forecasting. The hypothesis used indicates the information rate and energy consumption (Anand & Titus, 2017). There are several algorithms used for statistical signal processing with different applications such as the bootstrap method (Zoubir & Iskandler, 2007). However in this work data obtained via simulations are analyzed in order to interpret the simulation results. The efficiency of INACS is validated using a time series algorithm and Generalized Linear Model (GLM). The significance of the analysis is, it simultaneously associates the dependent and independent variables. The number of transmissions and energy consumption is analyzed using time series analysis and GLM analysis respectively.

4.5.1 Evaluation of ARIMA for INACS

The purpose of this ARIMA is to forecast the information relating to regression on itself to the series (number of transmissions versus time period). The first step is identification checking whether the series is stationary or not stationary. In the identification step if the series is stationary the process can be found with Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF). If the data is not stationary the differencing will make the non-stationary data to become

stationary. ARIMA (p, d, q) represents the number of lags of dependant variable. P indicates the number of transmissions. Number of times the variable is differentiated is denoted by (d), number of lags in error term is denoted by (q).

Null hypothesis: There is no relationship between the time duration and the reduction in the number of transmissions.

Alternate hypothesis: There is a relationship between the time duration and the reduction in the number of transmissions.

Dependant variable: Number of transmissions

Independent variable: Time period

The table 4.8 shows the values taken from simulation results for data analysis using SPSS (Statistical Package for Social Sciences). Table 4.9 denotes the appropriate model for Auto Regressive Integrated Moving Average. Table 4.10 shows the model fit statistics. The optimal order of differencing can be found with Root Mean Square Error (RMSE). Diagnostic checking can be done with Bayesian Information criteria (BIC), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Maximum Absolute Percentage Error (MaxAPE), Maximum Absolute Error (MaxAE).

Table 4.8: Time Period versus Number of Transmissions

Time duration in seconds INACS	Number of Transmissions INACS
10000	179000
50000	902000
100000	1079000

Table 4.9: Model description for numerical evaluations of INACS

Model Description		
Model id	Number of transmissions	Model Type ARIMA (0,0,0)

Table 4.10: Model fit statistics for Number of Transmissions in INACS

Number of Transmissions Model	RMSE	MAPE	MAE	Max-APE	MaxAE	Normalized BIC
	263187.414	33.056	142967.213	6.558	214450.820	25.694

Table 4.11: Residual ACF summary of INACS

Lag	Mean
Lag 1	-.664
Lag 2	0.164

Table 4.12: Residual PACF Summary of INACS

Lag	Mean
Lag 1	-.664
Lag 2	0.495

Residual value with ACF and PACF summary states the model fits is appropriate with white noise. The Auto Correlation Function (ACF) denotes the simple correlation between the current observation and the observation "p" period from current observation. It is shown in table 4.11. The Partial Auto Correlation Function (PACF) denotes the degree of association between the current observation and the observation "p" period from current observation. It is shown in table 4.12. The number of lags considered when developing a model can be demonstrated using the Residual ACF as in table 4.13 or Residual PACF as in table 4.14. Standard Error (SE) happening across the lag is also shown in Residual ACF and Residual PACF.

Table 4.13: Residual Auto Correlation Function for two lags in INACS

		Lag 1	Lag 2
Number of transmissions_Model_1	**ACF**	-.664	0.164
	SE	0.577	0.792

Table 4.14: Residual Partial Auto Correlation Function for two lags in INACS

		Lag 1	Lag 2
Number of transmissions_Model_1	**ACF**	-.664	-.495
	SE	0.577	0.577

Table 4.15: Predicted value of transmission using ARIMA model for INACS.

Time duration in seconds	Number of Transmissions	Predicted Number of transmissions
10000	179000	298139.34
50000	902000	687549.18
100000	1079000	1174311.48

The SPSS predicts the number of transmissions with the base line model. The actual number of transmissions are compared with these predicted number of transmissions as shown in table 4.15. Table 4.15 explains the actual number of transmissions and predicted number of transmissions for each assigned time duration using ARIMA model for INACS.

4.5.2 Evaluation of ARIMA for CWS

The number of transmissions incurred for CWS within the time period is shown in table 4.16. Table 4.17 denotes the appropriate model with the data obtained for Auto Regressive Integrated Moving Average (ARIMA). The goodness of fit measure can be found using table 4.18.

Table 4.16: Time Period versus Number of Transmissions

Time duration in seconds CWS	Number of Transmissions CWS
10000	476000
50000	1467000
100000	2192000

Table 4.17: Model description for numerical evaluations of CWS

Model Description		
Model id	Number of transmissions	Model Type ARIMA (0,0,0)

Table 4.18: Model fit statistics for Number of Transmissions in CWS

Number of Transmissions Model	RMSE	MAPE	MAE	Max-APE	MaxAE	Normalized BIC
	186050.958	10.367	101065.574	17.694	151598.361	25.000

Table 4.19: Residual ACF summary of CWS

Lag	Mean
Lag 1	-.664
Lag 2	0.164

Table 4.20: Residual PACF Summary of CWS

Lag	Mean
Lag 1	-.664
Lag 2	0.495

Table 4.21: Residual Autocorrelation Function for two lags in CWS

		Lag 1	Lag 2
Number of transmissions_Model_1	ACF	-.664	0.164
	SE	0.577	0.792

Table 4.22: Residual Partial Autocorrelation Function for two lags in CWS

		Lag 1	Lag 2
Number of transmissions_Model_1	ACF	-.664	-.495
	SE	0.577	0.577

Table 4.23: Predicted value of transmission using ARIMA model for CWS

Time duration in seconds	Number of Transmissions	Predicted Number of transmissions
10000	476000	560221.31
50000	1467000	1315401.64
100000	2192000	2259377.05

Table 4.19 shows the percentile of ACF of the residual with all the estimated models. Table 4.20 shows the percentile of PACF of the residual with all the estimated models. The residual autocorrelation function upon two lags is shown in table 4.21. Similarly the residual partial autocorrelation function for two lags is shown in table 4.22. From the numerical evaluations it is clearly evident where INACS yields a correlation indicating, alternate hypothesis is proved. The number of transmissions is reduced in compressive sensing with INACS than with CWS. We can see the reduced predicted number of transmissions for INACS in table 4.15 is better than the predicted number of transmissions of CWS as in table 4.23.

4.5.3 Generalized Linear Model for INACS

Null hypothesis: There is no significant difference between number of nodes and energy consumption.

Alternate hypothesis: There is a significant difference between number of nodes and energy consumption.

When the number of nodes within a sensing field is less the protocol consumes less energy. If scalability of sensor nodes is increased then the impact of energy consumption also increases. To interpret the scalability and energy consumption Generalized Linear model (GLM) is used. The numerical values obtained with MATLAB simulations in table 4.24 are used in GLM. EC_INACS: Energy consumption for Intelligent Neighbor Aided Compressive sensing. EC_CWS: Energy consumption for Compressive Wireless Sensing. The model information used is given in table 4.25 and the probability distribution in this case is normal with central mean. The link function is identity where dependant variable is not transformed and link can be used with any distribution.

Table 4.24: Number of Nodes versus Energy Consumption

Number of Nodes	Energy consumption INACS	Energy consumption CWS
10	4.02	4.31
20	5.47	6.32
30	6.72	8.26
40	8.72	9.97
50	10.32	12.02

Table 4.25: Model information of EC_INACS

Dependant variable	EC_INACS
Probability distribution	Normal
Link Function	Identity

The case processing summary is given in table 4.26 and it indicates the number of cases used. Table 4.27 shows the continuous variable information in order to check for no outlier in the data. Table 4.28 shows the deviation and scaled deviation for data values of EC_INACS and the Intercept model. Smaller value indicates better fit of AIC, BIC, CAIC. EC_INACS denotes the dependant variable. Intercept model is used. "a" represents the information criteria in smaller better form. "b" is the full log likelihood function displayed and it is used in computing information criteria. In table 4.29 the analysis type used for model effect is type III, generally applicable and does not need predictor variable. The dependant variable is EC-INACS with intercept model.

Table 4.26: Case processing summary of EC_INACS

Total number of variables in case processing summary	5

Table 4.27: Continuous variable information of EC_INACS

Variable	Value
N	5
Minimum	4.02
Maximum	10.32
Mean	7.0500
Standard deviation	2.51336

Table 4.28: Goodness of fit[a] of EC_INACS

Parameter	Value	df	Value/df
Deviance	25.268	4	6.317
Scaled Deviance	5.000	4	-
Pearson Chi-Square	25.268	4	6.317
Scaled Pearson Chi-Square	5.000	4	-
Log Likelihood[b]	-11.145	-	-
Akaike's Information Criterion (AIC)	26.290	-	-
Finite Sample Corrected AIC (AICC)	32.290	-	-
Bayesian Information Criterion (BIC)	25.509	-	-
Consistent AIC (CAIC)	27.509	-	-

Table 4.29: Tests of model effects of EC_INACS

	Type III		
Source	**Wald Chi-Square**	**df**	**sig**
Intercept	49.175	1	.000

Table 4.30: Parameter estimates of EC_INACS

Parameter	B	SE	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald-Chi-Square	df	Sig.		Lower	Upper
Intercept	7.050	1.0053	5.080	9.020	49.175	1	.00	1152.859	160.703	8270.427
Scale	5.054[a]	3.1962	1.463	17.456	-	-	-	-	-	-

In table 4.30 the test statistics of parameter estimates are shown. Fixed effects are observed from Wald test Confident Interval (CI) considering the parameters of covariance matrix. The dependant variable used in EC_INACS. “a” represents the maximum likelihood estimate

4.5.4 Generalized Linear Model for CWS

The model information used has been given in table 4.31. The link function is identity. In this model the dependant variable is not transformed and link can be used with any distribution. The distribution in this case is normal with central mean. The case processing summary has been given in table 4.32 indicates the number of cases used. Continuous variable information has been given in table 4.33 in order to check for no outlier present in data.

Table 4.31: Model information of EC_CWS

Dependant variable	EC_CWS
Probability distribution	Normal
Link Function	Identity

Table 4.32: Case processing summary of EC_CWS

Total number of variables in case processing summary	5

Table 4.33: Continuous variable information of EC_CWS

Variable	Value
N	5
Minimum	4.31
Maximum	12.02
Mean	8.1760
Standard deviation	3.01631

Table 4.34: Goodness of fit[a] of EC_CWS

Parameter	Values	df	Value/df
Deviance	36.393	4	9.098
Scaled Deviance	5.000	4	-
Pearson Chi-Square	36.393	4	9.098
Scaled Pearson Chi-Square	5.000	4	-
Log Likelihood[b]	-12.057	-	-
Akaike's Information Criterion (AIC)	28.114	-	-
Finite Sample Corrected AIC (AICC)	34.114	-	-
Bayesian Information Criterion (BIC)	27.333	-	-
Consistent AIC (CAIC)	29.333	-	-

In table 4.34 goodness of fit is shown. The smaller values indicate better fit in AIC, AICC, BIC and CAIC. The data interpretation using AIC emphasis on good predictions when the type of error is likely to be over fitting. BIC and CAIC can be used when the type of error likely to be under fitting. “a” represents the information criteria are in smaller better form. “b” is the full log likelihood function displayed and it is used in computing information criteria. The dependant variable is EC_CWS with model intercept.

In table 4.35 the analysis type used for model effect is type III. It is generally applicable and does not need predictor variable. Type III sum of squares is used as the default model to evaluate hypothesis.

Table 4.35: Tests of model effects of EC_CWS

	Type III		
Source	**Wald Chi-Square**	**df**	**sig**
Intercept	45.921	1	.000

Table 4.36: Parameter estimates of EC_CWS

Parameter	B	SE	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald-Chi-Square	df	Sig.		Lower	Upper
Intercept	8.176	1.2065	5.811	10.541	45.921	1	.00	3554.608	334.038	37825.702
Scale	7.279[a]	4.6033	2.107	25.141	-	-	-	-	-	-

In table 4.36 the corresponding test statistics with its confident interval along with exponential parameters are shown. The dependant variable used is EC_INACS. "a" represents the maximum likelihood estimate. In both the proposed protocols alternate hypothesis is proved. It shows the strong correlation with number of nodes and energy consumption.

4.6 Conclusion

INACS focus on achieving reduction in sensing and transmission cost. The ability to capture the characteristics of sensor data and design of measurement matrix mapping using the spatial and temporal coordinates is established in INACS. Recovery performance is facilitated by convex optimization. In addition INACS emphasize on data forwarding. The Rademacher matrix is formed using Pearson's correlation coefficient and this matrix is used to analyze the correlation between nodes for data forwarding. The data forwarding from source node to one of the neighbor nodes happen based on highest correlation. The proposed data-driven approach based on compression at the source, data forwarding and recovery at the destination achieves a greater reduction in the number of transmissions than CWS. The significant results of INACS is achieved by load balancing features of internetworking indicating superiority in terms of reduction in the number of transmissions, energy efficiency and reconstruction accuracy. The scalability of the INACS algorithm is analyzed by varying the metrics such as transmission radius, number of nodes and time duration. The number of transmissions and the energy consumed is studied based on these metrics. Time series analysis has been used for characterizing the multivariate behavioral pattern between the number of transmissions and the time period. In addition the forecast is also made on energy consumption and the scalability of nodes within the sensing field revealing the strong relationship between both the metrics using Generalized Linear Model (GLM).

CHAPTER 5: COMPRESSIVE SENSING WITH PERCEPTRON BASED FORWARDING

5.1 Introduction

Compressive sensing with data acquisition in wireless sensor networks has overcome the rigid sampling procedure to acquire sensor data. In the wireless network the efficiency of data transmission is considerably reduced due to congested links. In order to ensure efficient transmission from source to sink through intermediate forwarding nodes a desired data forwarding threshold has to be established. Initially the physical data has to be made sparse by using Discrete Wavelet Transform. The data forwarding threshold needs to be estimated for the sparse physical data when it begins its transmission. This process of assigning threshold values for data forwarding would save the energy consumption of the overall network.

In the framework of INACS compressive sensing at source and reconstruction at sink has been performed. The co-located sensor performs sensing using spatial and temporal process thereby reducing redundancy in sensing and transmission. In INACS the transmission capacity of intermediate nodes within a wireless channel has not been analyzed. Analyzing the bandwidth resources, updating the status of wireless links is not performed in INACS. To resolve these issues machine learning techniques can be incorporated. A dynamic framework is proposed analyzing the bandwidth, traffic intensity or capacity of forwarding nodes for transmission. The framework accommodates the transmission from source to sink via intermediate forwarding nodes depending on the desired data forwarding threshold. This process has been incorporated by using Single Layer perceptron considering two scenarios. For scarce network resources communication is established when the target response (Threshold) exactly matches between two source nodes and intermediate node.

The input values should neither be lesser nor exceed but should exactly match the target value. For networks of surplus bandwidth communication is established if any one of the source nodes associated with intermediate node just exceeds the data forwarding threshold.

5.2 Previous Work on Distributed Compressing Scheme

In a wireless environment the transfer of information plays a vital role in determining the available resources. An information-driven approach with compressive sensing is used to show the logical mapping of sensory data can conserve energy resources and bandwidth in multi-hop networks (Dang et al., 2007). In On-route compressive sensing, rather than compression at source and decompression at the destination intermediate relays or powerful nodes are deployed for compression (Razzaque et al., 2013). Similarly in Distributed Compressed Estimate (DCE) scheme rather than considering the entire time duration and compressing the observed data intermediate nodes could exchange the measurement matrix with a local estimator (Xu et al., 2015). Thus DCE requires minimal bandwidth and achieves a low mean squared error (MSE). However, spatially correlated sensors exhibit similar sparsity and a similar measurement matrix. Bottleneck link in wireless networks occurs due to varying rate of data traffic. It can be measured and estimated using machine learning techniques. Recent work on compressive sensing deal with proper utilization of bandwidth. In addition it also describes a deep neural network and its significance in a wireless environment. The feature extraction layer used in this communication framework characterizes the bandwidth, noise and fading characteristics for proper recovery (Lu & Bo, 2019).

The foremost aspect of compressive sensing is sparsity. A signal must be sparse in nature and if it is not, a transformation must be performed to ensure it exhibits sparsity. A sparse representation of wireless signals with the desired sparsity level is predominantly achieved using pursuit methods, dictionary-based methods and channel estimation via sensing properties (Qin et al., 2018). In this proposed work a discrete wavelet transforms (DWT) is used to ensure sparsity. The significance of DWT is it can match the distributed nature of data aggregation and intermittent connectivity occurring

during data transmission in wireless sensor communication networks (Acimovic et al., 2005). Hence discrete data can be recovered through an appropriate transformation using the wavelet coefficient. The wavelet transform is well suited for deterministic signals. The deterministic signals are correlated by nature. Therefore when there is de-correlation of signal involved in data-gathering, design of the wavelet coefficients should match the number of discontinuities (Acimovic et al., 2005). The discrete wavelet transform has been used with a filter bank in one-dimensional wireless sensor networks. When there is a subsequent increase in the number of levels of decomposition, the local information cost (overhead) between sensors increases (Ciancio et al., 2006). Hence the proposed algorithm attempts to compensate for overhead in addition to enabling the desired traffic using perceptron-based forwarding.

Multi-scale analysis has been done with graph wavelets (Crovella & Kolaczyk, 2003). It can quantify the network traffic with a decision criterion for analyzing the traffic intensity. The discussion reveals the impact of network topology and traffic characteristics on subsequent data from the sensor nodes. Wavelet transform is suitable for irregular structure termed as "point-like singularities" as in with practical sensor network (Guo et al., 2009).

5.2.1 Machine Learning Techniques in Wireless Sensor Networks

Computational intelligence in wireless sensor networks refers to the property of possessing the capability of inputting raw sensor data and producing reliable-timely responses with minimal fault tolerance (Kulkarni et al., 2011). Incorporating machine learning in wireless sensor networks has been surveyed in terms of both functional and non-functional aspects (Alsheikh et al., 2014). Functional aspects include routing and analyzing the data link layer operations. Non-functional aspects include security issues and the integrity of data (Alsheikh et al., 2014). A Markov decision process can be incorporated to achieve better performance although this depends on the application.

The applications discussed with regard to wireless sensor nodes are concerned with optimal transmission levels, selection of the intermediate relay from the source and data aggregation techniques (Alsheikh et al., 2015).

Multi-dimensional sensing matrix has been proposed along with multi-linear dictionary for sparsifying with "Tensor based compressive sensing" (TCS) (Ding et al., 2017). Multi-dimension projections are made by segregating it into separable and non-separable entities. This approach using TCS is better than the random sensing matrix. The limitation of random sensing matrix is it imposes computational complexity at recovery due to unstructured nature (Candes & Romberg, 2007). The survey of literature on wireless sensor networks with machine learning reveals, 67% of articles use supervised learning approaches (Kumar et al., 2019). Machine learning does not produce instant results; rather it takes time to learn from historical data consuming more resources in a wireless environment. Thus the incorporation of machine learning should complement the reporting interval at the sink (Kumar et al., 2019). In such scenarios a hop-by-hop pattern to achieve the desired data rate with minimal distortion in artificial neural networks has been used (Alsheikh et al., 2016).

Wireless sensor node consists of sensing unit, processing unit and power unit. In the sensing unit the analog signal is converted to digital form. This conversion process is performed below sampling rate to reduce the number of samples. The processing unit interfaced with the transceiver transmits the sparse data and can be recovered at sink. The data flow between the sensor and sink is modelled using perceptron logic and its truth table. The perceptron consists of single layer of input from each source to an intermediate node. The output layer of intermediate node is determined by the source node and its weight values. W_1 and W_2 are the weight values of input 1 and input 2 respectively as shown in figure 5.1.

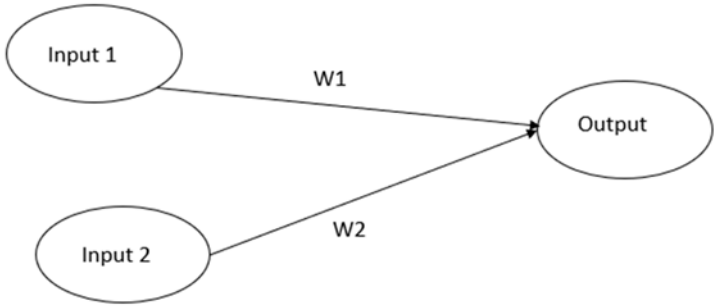

Figure 5.1: Architecture of single layer perceptron

Figure 5.1 depicts the architecture of single layer perceptron. Transmission is performed based on the weight values, analogous to the link in a wireless communication. Target threshold is assigned priorly at intermediate nodes and if target threshold is not met by the forwarding nodes then subsequent weight values are being altered.

5.2.2 Topological Significance of Compressive Sensing

The selection of projection nodes (relay nodes) in a random fashion can undermine the performance of compressive sensing by increasing the number of transmissions (Qiao & Zhang, 2018). Centrality nodes play a major role in transfer of data in any communication network. Hence one hop information, localization or analyzing the centrality measure has to be known for analyzing the network (Mahyar et al., 2018). Hierarchical adaptive spatial-temporal data compression is proposed incorporating spatial and temporal compressions. The observed sensory data are transformed using the discrete cosine transform where the obtained transform vectors are based on the threshold values related to location coordinates in the sensing field. The cluster head exploits temporal information using the obtained readings for suitable compression and transmission via two-dimensional discrete wavelet transforms (Chen et al., 2019). Layered adaptive compressive design for efficient data gathering (LACD-EDC) is proposed in (Chen et al., 2019).

The cluster head exploits spatial correlation with intra-cluster and inter-cluster communication. LACD-EDC uses an adaptive dictionary and can process multi-resolution data achieving significantly better performance than the predefined dictionary used for sparsifying databases (Chen et al., 2019). However the process of selecting a cluster head or a superior node in transmission should not increase the compression ratio and burden of wireless links. Hence the proposed network model in this work uses two nodes associated with a relay to forward compressed data. Topology control has been proposed wherein the network is being portioned into multiple strips. The strip center is responsible for compressive sensing and routing to the sink (Li et al., 2018). Random walk based compressive sensing (RWCS) has been proposed to deduce the cost in transmission and measurement using ring topology. In RWCS a measurement matrix is created based on compensation function (distance to the sink) (Zhang et al., 2018).

"Link Estimation with Sparse Sampling" (LESS) has been proposed with topological and routing changes along with link quality through reduced overhead using compressive sensing (Zhou et al., 2018). The tribulation behind the LESS protocol is entirely dependent on the sink. It decides compression factor and routing path. Deep learning approach with graph wavelet has been used for compressive sensing and routing in (Li et al., b, 2018). With minimal training set the maximum magnitude of data acquisition and transfer can be obtained. Similarly "Light Weight Compressed Data Aggregation" (LWCDA) is proposed for non-overlapping. The compression rate of LWCDA is better and cost of transmission is also less in (Amarlingam et al., 2018).

5.2.3 Traffic Intensity in WSN

In order to effectively transfer the compressed data there should be minimal number of sensors needed to ensure connectivity to the sink. This is termed as communication constraint. The cardinality constraint should also be satisfied. The active status of a node with its timeslot allocated for transmission should be greater than the

required number of timeslots to execute compressive sensing (Du et al., 2018). The classification of traffic intensity has been roughly divided into fixed (predetermined path) or Markov (random path) classes with maximum likelihood estimation (Vardi 1996). However in the scenario where the number of samples increases the bias factor towards certain nodes causes congestion. Therefore fixed path routing reduces the quality of transmissions during data transfer between source and destination due to congestion.

The sparsity-regularized matrix formulation (SRMF) with a compressive sensing approach classifies high-rate traffic and low-rate traffic. A deep belief network is used for classification of low-rate traffic. Spatial-temporal compressive sensing is used for high-rate traffic conditions. The study shows appropriate classification of traffic condition reducing the problem of bottleneck bandwidth (Nie et al., 2018). If the number of measurement vectors required, matches the sparsity conditions then it subsequently burdens the recovery process at the sink. Hence a mixed support model is used where the underlying joint sparsity is estimated for obtaining the measurement vector. This process is performed by cluster averaging the value of information with a suitable measurement vector using Eigen value decomposition (EVD) (Liu et al., 2018).

The measurement matrix and the sparsity matrix must be incoherent in Compressive Sensing (CS). Determined by a known chaos sequence in (Alwan & Hussain, 2019). The result shows the known chaos sequence has higher reconstruction accuracy than conventional compressive sensing, because the control parameters of the chaos sequence are predetermined by the sink (Shekaramiz et al., 2019). Sparse Bayesian learning (SBL) has been used for multiple measurement vectors (MMV). The sparsity of the signal is achieved using an unknown binary support learning vector. In addition a total variation support vector for sparsity is obtained via the sum or difference of the signals (Shekaramiz et al., 2019). However the unknown signal rarely exhibits the

same non-zero entities in an MMV. Two methods of compressive sensing based on channel identification have been used at two different sparsity levels (Van Nguyen et al., 2018). The first uses blind compressive sensing identifying the signal-to-noise ratio (SNR) metric and assigning the appropriate sparsity level. The second assigns cluster-level sparsity when the channel characteristics are previously known (Van Nguyen et al., 2018). An optimal compression matrix has been obtained in homogeneous sensors where the noise covariance can be scaled identically to attain a minimal bound of distortion (Zhang et al., 2019). Determining the noise and anisotropic conditions of the channel makes it difficult to obtain the parameters of compressive sensing in large-scale sensor networks. Expander graphs have been used for identifying the sparsity matrix using delay vectors (Firooz & Roy, 2014). However expander graphs are not suitable for delay-sensitive applications and identifying the network conditions is challenging.

This section of traffic analysis deals with internetworking. Mixed integer nonlinear programming (MINLP) examines the traffic rate at the upper bound using nonlinear programming (NLP) with the help of rank functions. The property of the channel in the case of MINLP is a single-step Markov process with Eigen decomposition (Zhou et al., 2015). Load balancing has been achieved with stochastic routing to eliminate stale routes and to deliver data to the sink. The Markov chain specifies the error bound for recovery of compressed data with minimal transmission cost (Huang & Soong, 2019). The compressed data must traverse a proper route for proper recovery in wireless sensor networks.

Impact of forwarding nodes “with” and “without cache” using decode and forward techniques has been discussed (Xia et al., 2018). It states the content delivery diversity and signal strength variations in wireless environment. If relay nodes are positioned with cache then the destination can recover the data from relay rather than from the source. The discussions of production traffic and consumption traffic with and

without data aggregation has been discussed in (Ma et al., 2018). The discussion states if the wireless channel is occupied by more packets then subsequently network efficiency is reduced in the period.

5.2.4 Energy Impact of Compressive Sensing Algorithms

Wireless sensor nodes with limited energy and a limited communication range usually exchange information with nearby nodes. By decreasing the volume of transmissions minimal depletion of energy can be achieved (Mahfoudh & Minet, 2008). In "Compressed Network Coding based Distributed-data Storage" (CNCDS) the data transmission and reception consumes almost the same energy for short distance. Thus the probability of forwarding has to be adjusted according to the capability of intermediate neighbors (Yang et al., 2013). Covariogram-based Compressive Sensing (CBCS) was proposed where the underlying signal statistics and correlation are determined by online estimation. The energy model used can be classified into the following: the energy consumption used in compression, the second is energy consumption involved in transmission of information from sensor to sink and energy cost used for exchange of control packets (Hooshmand et al., 2015).

Two-dimensional data-gathering is proposed using a discrete cosine transform (DCT) and a discrete Fourier transform (Dolas & Ghosh, 2018). The increasing measurements with minimal sparsity results in attaining higher recovery with minimum energy consumption and negligible RMSE (Dolas & Ghosh, 2018). Rotating random sparse sampling proves the minimal energy consumption achieved by changing sensor nodes into sleep mode. The transfer of data in the remaining awake nodes occurs in a rotational manner with diverse sampling ratios (Xu et al., 2019). Integrated data and Energy gathering protocol (iDEG) was proposed with single hop communication (Jain et al., 2019). The iDEG protocol implement "partial canonical matrix" for reducing sensing time. In addition it balances energy inefficiency using energy harvesting at

specific nodes. The discussion of balancing energy efficiency and recovery is done with temporal and spatial redundancy. The process of reducing temporal redundancy has been done considering different clusters of wireless sensor network. Spatial redundancy is obtained with hybrid compressive sensing to attain the minimal projection for reconstruction (Zhou et al., 2019).

5.3 The Framework - Perceptron-based Optimal Routing (POR) and Perceptron-based Routing with Moderate Traffic Intensity (PRMTI)

The motivation behind this work is development of routing protocols considering the network resources. Perceptron based Optimal Routing (POR), Perceptron based Routing with Moderate Traffic Intensity (PRMTI) is proposed. POR and PRMTI transmission protocol can be used for scarce and surplus bandwidths respectively. POR consumes lesser energy, decreases transmission cost but the convergence time is delayed. PRMTI shows increased energy consumption, increases transmission cost and convergence time is quicker. The objective of the proposed framework is to perform compressive sensing using a discrete wavelet transform and to transfer data with the aid of perceptron. The proposed protocols are designed in a network model with a decision criterion for forwarding where two nodes are associated with a relay node. In this work the following two assumptions are made. Compressive sensing is performed, where sparse vectors are quantified on an absolute scale. The decision to forward or not forward is translated into a linear separable problem based upon the network resources. The topology is designed with a sink at the center. Data transfer occurs from a pair of sensors through the relay to the sink.

The physical data is compressed and our work focuses on data forwarding from source to intermediate forwarder after estimating the network resources. The intermediate forwarder can also be termed as relay node. The method of data forwarding is formulated into two perceptron techniques depending on bandwidth or network resources. The first criterion: POR is proposed when the network resources are not abundant. It uses a bipolar logic and a target value at the relay node for data forwarding. If the target response at the relay node matches with the input values of the source nodes then data is transmitted from the source node to the relay node. If there is a mismatch then the weight updates are performed iteratively until the incoming data

meets the target response and data is forwarded after the iterations. The second criterion: PRMTI is proposed when the network resources are abundant. PRMTI forwards data from source node to the relay node when the link capacity exceeds the defined data-forwarding threshold at the relay node. If the input value at one of the sources does not exceed the target or threshold value then the weights and error calculations are continuously updated until the condition of PRMTI is satisfied for data forwarding.

Assuming, the two source nodes forward data to the relay node. The POR algorithm transmits data only when it finds the exact match between the two source nodes and target value at relay node. The input values and target values should not be lesser nor higher but must be equal for data forwarding to take place. But in PRMTI the relay accepts the data even if any one of the source nodes exceeds the target value. It means the input value at source and target value at the relay need not be exactly same. Once the input values exceed above threshold (target value) the data can be transmitted. Hence any one of the algorithms is used for data forwarding to the relay node based on the availability of network resources. The final process is signal recovery obtained by using in-coherence property of sensing and observation matrix. The signal generation, data forwarding logics of POR and PRMTI, model calculations and signal recovery are explained in the following sections. The simulations results show that the number of transmissions to be higher in PRMTI than compared to POR. Data analysis is also performed on the simulation results to validate the results of POR and PRMTI. Figure 5.2 explains the framework of POR and PRMTI. The network is initialized and configured.

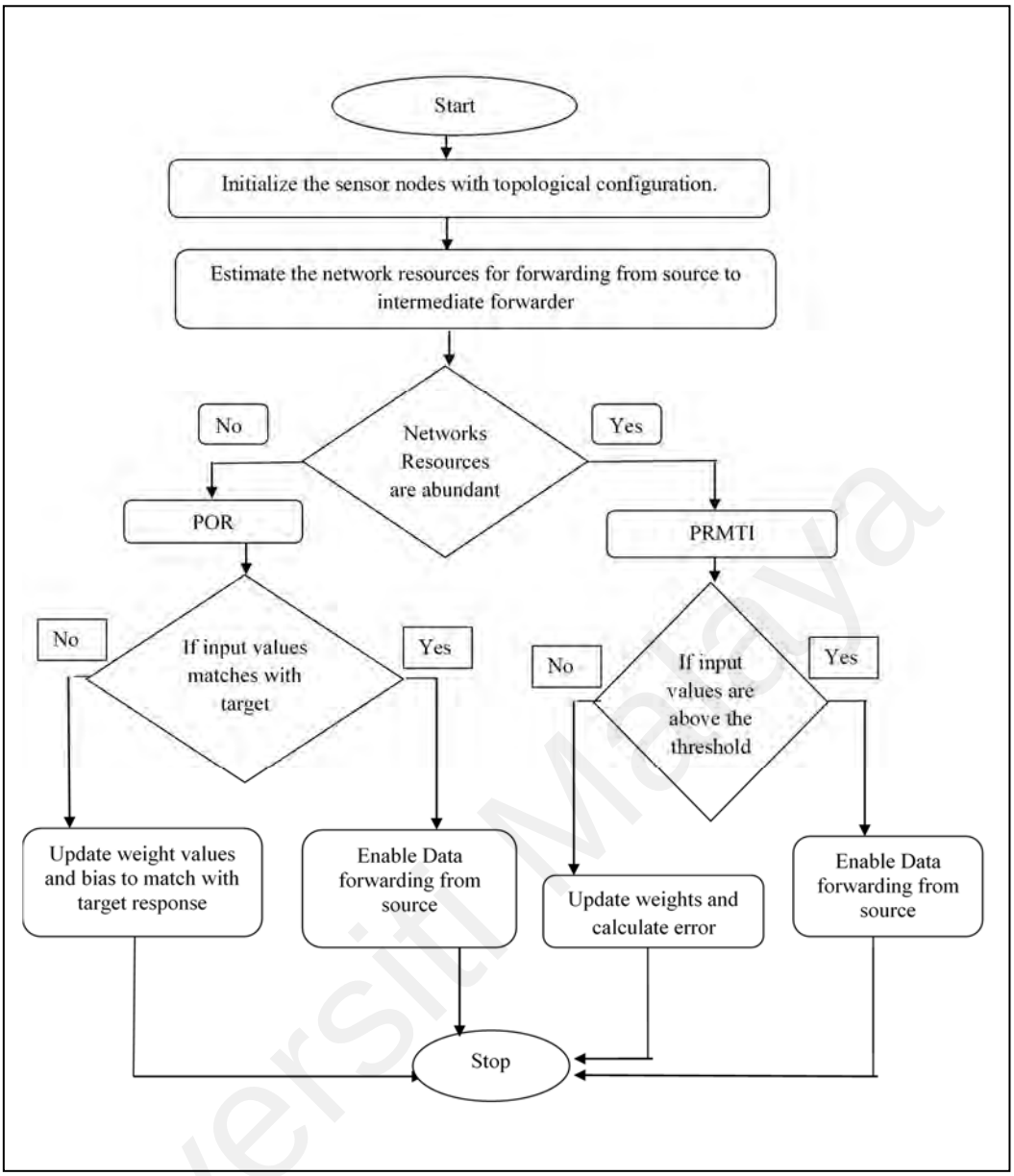

Figure 5.2: Illustration of POR and PRMT framework

5.3.1 Generation of Signal

Let $\bar{x}$ be the unknown data procured from sensor network represented by

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_N \end{bmatrix} \tag{5.1}$$

N denotes an N-dimensional signal vector. M denotes the number of observations. Then $\bar{y}$ denotes observation matrix. The sensing matrix is denoted byϕ_s.

$$\bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_M \end{bmatrix} \tag{5.2}$$

$$\phi_s = \begin{bmatrix} \phi_1^T \\ \phi_2^T \\ \phi_M^T \end{bmatrix} \quad (5.3)$$

The recovery of unknown data can be obtained from equation 5.4.

$$\bar{y} = \phi_s \bar{x} \quad (5.4)$$

For any data obtained y_i is the projection of x_i on sensing matrix ϕ_{si}. The number of observations M should be greater than or equal to N. The following process after signal generation is data forwarding where POR or PRMTI algorithm would be used based on network resources. The two algorithms are explained in detail as below.

5.3.2 POR - Data Forwarding

The development of Perceptron based Optimal Routing (POR) considers a scenario where network resources are scarce. Therefore POR protocol tries to minimize the cost of transmissions. An example scenario of a node based on weight updation with bipolar inputs is shown in figure 5.3. Its truth table is shown in table 5.1. r_{in} is the current status of forwarding inputs. A_1 and A_2 represent two nodes with weight values w_{1R} and w_{2R}. The binary inputs of forwarding and not forwarding is represented as a_1 and a_2 along with weight values w_{1R} andw_{2R}. The target value is denoted by t.

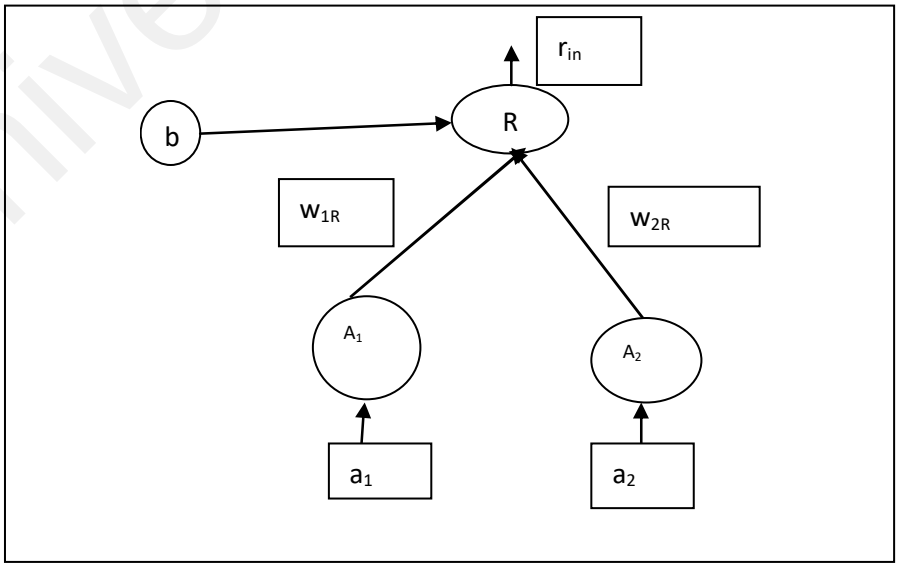

Figure 5.3: POR- Intermediate forwarding at relay node

Table 5.1: Permutations of routing with optimal traffic intensity-POR.

	a_1	a_2	t
Logic 1	1	1	1
Logic 2	1	-1	-1
Logic 3	-1	1	-1
Logic 4	-1	-1	-1

$$f(r_{in}) = \begin{cases} 1 & r_{in} \geq 0 \\ 0 & r_{in} = 0 \\ -1 & r_{in} \leq 0 \end{cases} \quad (5.5)$$

The probability of node forwarding is governed by an activation function as in equation 5.5. Initially $w_{1r} = 0$, $w_{2r} = 0$ and b=0. "b" denotes the bias weight. The first input permutation is taken for updating weights [1, 1, 1] as in logic 1. Substituting the values in equation 5.6.

$$\boldsymbol{r_{in}} = b + w_{1R}a_1 + w_{2R}a_2 \quad (5.6)$$

$$\boldsymbol{r_{in}} = 0 + 0 \times 1 + 0 \times 1$$

$$\boldsymbol{r_{in}} = 0$$

$$\boldsymbol{R} = f(r_{in}) = 0 \quad (5.7)$$

For forwarding equation 5.8 should be satisfied.

$$f(r_{in}) = t \quad (5.8)$$

$f(r_{in}) \neq t$in the first case hence forwarding does not take place.

In the next step updating of weights has to be done as in equation 5.9 to 5.11.

Updating weight for the first link between A_1 and R gives:

$$w_{1R(new)} = w_{1R(old)} + \alpha t a_1 \quad (5.9)$$

Updating weight for the link between A_2 and R gives:

$$w_{2R(new)} = w_{2(old)} + \alpha t a_2 \quad (5.10)$$

Where the value of α represents learning rate (α=1).

$$b_{(new)} = b_{(old)} + \alpha t \quad (5.11)$$

Updating and bias values changes the weights as $w_1 = 1, w_2 = 1$ and b =1. Logic 2 input patterns are [1, -1, -1]. Substituting these values in equation 5.6.

$$\boldsymbol{r_{in}} = 1 + 1 \times 1 + 1 \times -1$$

$$\boldsymbol{r_{in}} = 1$$

$$\boldsymbol{R} = f(r_{in}) = 1 \qquad (5.12)$$

$1 \neq -1 : f(r_{in}) \neq t$. Forwarding does not take place.

Updating weight and bias values changes the weights as $w_1 = 0, w_2 = 2$, b =0. The third input pattern is [-1, 1, -1]. Substituting the values in equation 5.6.

$$\boldsymbol{r_{in}} = 0 + 0 \times -1 + 2 \times 1$$

$$\boldsymbol{r_{in}} = \boldsymbol{2}$$

$$\boldsymbol{R} = f(r_{in}) = 1 \qquad (5.13)$$

$1 \neq -1: f(r_{in}) \neq t$. Forwarding does not take place.

Updating weight and bias values changes the weights as $w_1 = 1$, $w_2 = 1$, b = -1. The fourth input pattern is [-1, -1, -1]. Substituting the values in equation 5.6.

$$\boldsymbol{r_{in}} = -1 + 1 \times -1 + 1 \times -1$$

$$\boldsymbol{r_{in}} = -3$$

$$\boldsymbol{R} = f(r_{in}) = -1 \qquad (5.14)$$

$\boldsymbol{-1} = -1 : f(r_{in}) = t$. Forwarding takes place.

Hence for the first three permutations the data transmission is denied and the condition is satisfied only for the fourth permutation. In POR the forwarding nodes are named as a_1 and a_2. The capability of forwarding is determined by associated intermediate node with target value (t). The weight values are analogous to the link capacity and are initially set as zero to provide sufficient time for compressive sensing. Simultaneously if the target is not met by the forwarding nodes then updating weight values takes place. After substituting and updating the weights only logic 4 has a threshold value enabling forwarding. When the criterion for forwarding between two nodes and its relay matches the target output then the transmission is initiated using POR. The advantage of POR is, it is suitable in events where the network resources are scarce. However the reporting interval to the sink in POR is long.

In case of POR the weight values can alter the target function of nodes for forwarding the data. Initially the weight values are set to zero to provide sufficient time to compress the data. Then if the target response does not match the corresponding truth table new weights are calculated using equations 5.9 to 5.11. These new weights are applied to the next values of the corresponding columns in the truth table. If this matches the target response the corresponding forwarding occurs otherwise transmission does not occur.

5.3.3 PRMTI - Data Forwarding

The motivation behind the development of the PRMTI protocol is its quick convergence. Hence the reporting interval in the transfer of data to the sink is short. Consider the example scenario for data-forwarding shown in figure 5.4 and its truth table in table 5.2.

Table 5.2: Permutations of routing with optimal traffic intensity–PRMTI

	a_1	a_2	a_1 AND a_2
Logic 1	0	0	0
Logic 2	0	1	0
Logic 3	1	0	0
Logic 4	1	1	1

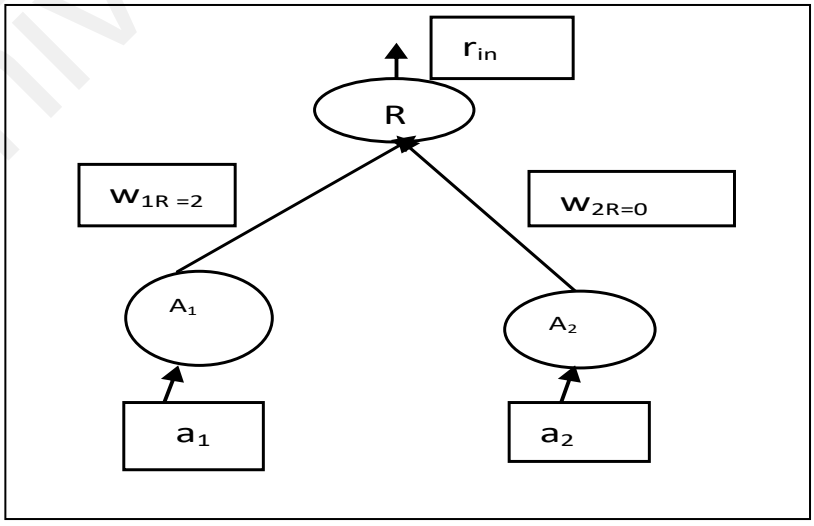

Figure 5.4: PRMTI - Intermediate forwarding at relay node

The threshold value of the forwarding node R is assigned in such a way, if it exceeds the data-forwarding value, then forwarding occurs.

$$R = f(r_{in}) = \begin{cases} 1.5 & r_{in} \geq 0 \\ 0 & r_{in} \leq 0 \end{cases} \quad (5.15)$$

$$\text{Incoming signal} = \text{Sum}(\text{node value} \times \text{weight of link}) \quad (5.16)$$

$$\text{Error}=(\text{Right value-Actual value})^2 \quad (5.17)$$

Range of error values are determined by the actual value. In all cases the error term should be less than the actual value being denoted as forwarding threshold. The weight is updated using equation 5.18.

$$W_{new} = W_{old} + (Input \times error \times learning\ rate) \quad (5.18)$$

The feed forward architecture of PRMTI is explained. Initially the weight values are w_{1R}=2 and w_{2R}=0. The output threshold for forwarding is R=1.5. Hence according to table 5.2 logic 3 and logic 4 provide forwarding since multiplying the weight gives a value exceeding the threshold of R. The PRMTI protocol does not use the criterion of matching the target response for forwarding. Hence if the relay node exceeds the forwarding threshold for any one of the associated senders it transmits data. The advantage of PRMTI is its quick convergence time. The PRMTI protocol consumes more energy due to subsequent forwarding. In the second case of PRMTI if any of the weight values exceed the activation threshold then forwarding criterion is satisfied. If the threshold value is lower, then forwarding occurs by updating the weights using equation 5.21 depending on the error terms learning rate and the input function. The entire topology of the feed-forward neural network is shown with an example in figure 5.3. The data initialization path and routing path from the source nodes are shown in figure 5.5. An example is also shown in figure 5.6 where the weight value of the link between node 1 and node 3 is w_{13}.Similarly the weight value of the link between node 2 and node 3 is w_{23}.The decision criterion of node 3 depends on node 1 and node 2.

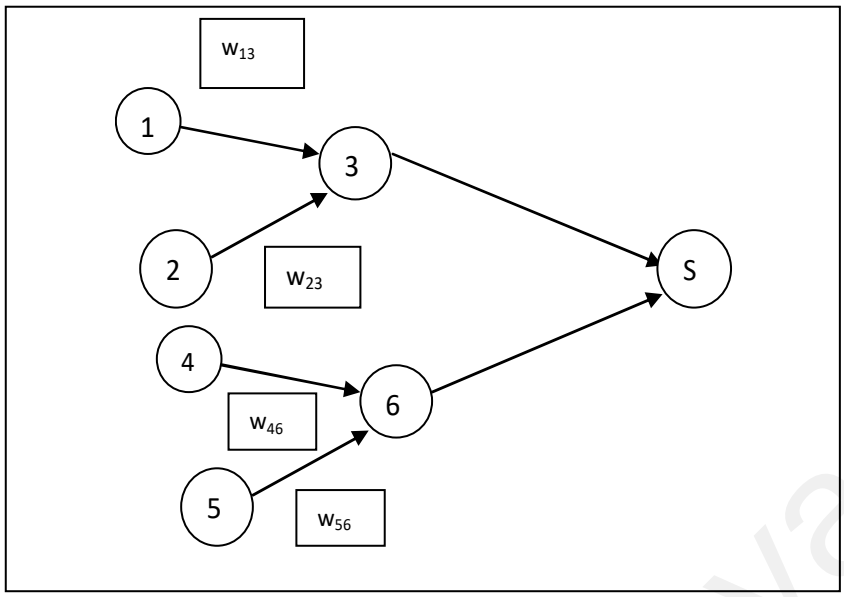

Figure 5.5: PRMTI - Fully connected feed-forward architecture

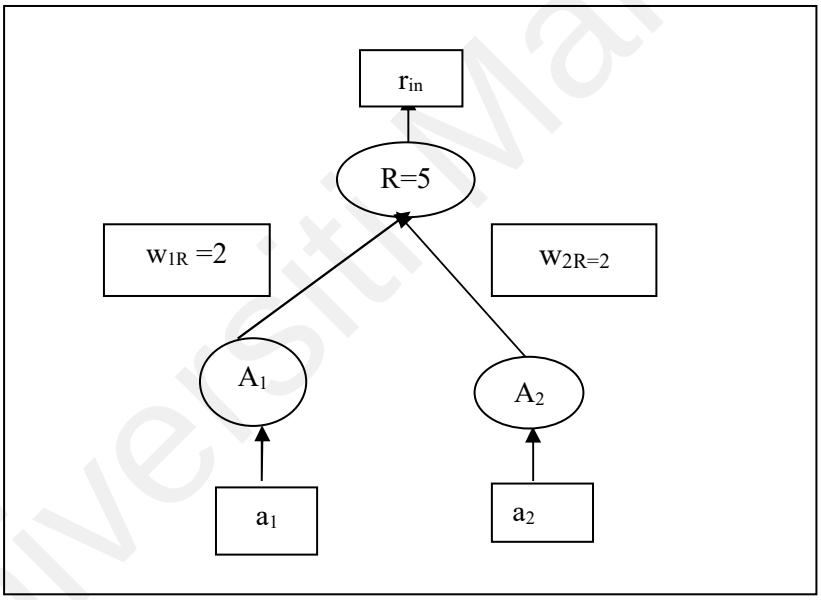

Figure 5.6: PRMTI - Illustration of intermediate forwarding at relay node

Similarly the weight value of the link between node 4 and node 6 is w_{46} and the weight value of the link between node 5 and node 6 is w_{56}. The decision criterion based on node 6 depends on node 4 and node 5. Finally node 3 associates itself with node 6 in forwarding to S. An example of intermediate forwarding in PRMTI is explained. Sensed data is made sparse with the help of a Discrete Wavelet Transform (DWT) and is ready for transmission from source. However, the condition of data flow to sink is based on two logics namely the forwarding capability of intermediate node and the link

availability from source node to intermediate node. In this case forwarding ability is determined by minimum processing capability (value) of the respective node and its associated metrics as in table 5.2.

Figure 5.6 depicts the illustration of intermediate forwarding. In this case we assume the minimum processing capability of intermediate forwarder as 5 or we can say the minimal weight value or forwarding capability of sensor is 5. Hence sensor 1 starts its transmission if the weight value exceeds 5 or it refrains from transmission. The weight assigned depends on link capacity and data rate generated. Similarly sensor 2 starts its transmission if the weight value exceeds 5 or it refrains from transmission. It can also be stated as when the combination of weights exceeds the threshold value at the source nodes.

$$R = f(r_{in}) = \begin{cases} 5 & r_{in} \geq 0 \\ 0 & r_{in} \leq 0 \end{cases} \qquad (5.19)$$

The weight is updated using equation 5.18. Initially the weight values of w_{1R}= 6, and weight vale of w_{2R}= 0. As per logic 3 in table 5.2 if the incoming relay node exceeds the forwarding threshold for any one of the associated senders it transmits data. In the second case the weight value of w_{1R}= 2 and weight vale of w_{2R}= 2, a_1=1, a_2=0. In this scenario the weight values are less than the data forwarding threshold at the relay node hence updating of weight takes place as below:

Updating weight $w_{1r:}$

$$\text{Incoming signal} = (1 \times 2)$$

$$\text{Incoming signal} = 2$$

$$\text{Error}=(\text{Right value-Actual value})^2 \qquad (5.20)$$

$$\text{Error}=(5\text{-}2)^2$$

$$\text{Error}=9$$

$$W_{1new} = W_{1old} + (Input \times error \times learning\ rate)$$

$$W_{1new} = 2 + (1 \times 9 \times 0.5)$$

$$W_{1new} = 6.5 \quad (5.21)$$

Updating weight w_{2r}can be done as below.

$$\text{Incoming signal} = (0 \times 2)$$

$$\text{Incoming signal} = 0$$

$$W_{2new} = 0 \quad (5.22)$$

Hence forwarding happens from node a_1 to the intermediate node R after updating of weight W_{1r}.The weight of W_{2r} is not updated as the incoming signal does not satisfy the data forwarding threshold value. The weight values for the subsequent iteration will provide faster transmission until the intermediate node is susceptible to transmit. Once the intermediate node is unable to transfer it refrains of transmissions from senders changing its R value. At source Discrete Wavelet Transform is performed and recovery at sink takes place using Inverse Discrete Wavelet Transform.

5.3.4 Reconstruction Phase

Sparsity can be enforced according to equation 5.23.

$$min\|x\|_0 \quad (5.23)$$

Maximal incoherence should occur between the sensing matrix and sparsity basis.

$$\bar{x} = \psi_s \theta \quad (5.24)$$

Whereψ_s is a basis matrix.θ denotes non-zero entries.

Thus the equation can be rewritten as in Equation 5.25.

$$\bar{y} = \phi_s \psi_s \theta \quad (5.25)$$

Thus the process of signal compression, data forwarding using POR and PRMTI, signal recovery is explained in this section 5.3.

5.4 Results and Discussion

In both proposed algorithms compression is performed using discrete wavelet transform and transmissions from source leverage on communication resources. The machine learning algorithm with perceptron-based data transfer works by associating the sensors with one another based on available resources. Distributed computing in wireless sensor network occurs at the wireless interface determining the choice between forwarding or not based on available bandwidth. Hence the communication protocol developed should meet the requirements of communication bandwidth which is scarce long and in-coherent with an enormous data rate produced by the sensors (Yick et al., 2008). Initial weight values are assigned based on the coverage and location coordinates of node to its forwarders. Data forwarding threshold values in POR is assigned based on carrier sensing inputs matching the target values.

Data forwarding threshold values in PRMTI is assigned in such way, the carrier sensing inputs is greater than forwarding threshold at intermediate nodes. In this work the following two assumptions are made. Compressive sensing is performed where sparse vectors are quantified on an absolute scale. Then the decision to forward or not forward is translated into a linear separable problem based upon the network resources. The topology is designed with a sink at the center. Data transfer occurs from a pair of sensors through the relay to the sink. Simulation parameters are shown in table 5.3.

Table 5.3: Simulation parameters of POR, PRMTI and CDG

Parameters	Values
Number of sensor nodes	50
Transmission radius	100 m
Initial energy of sensor nodes	2 J
Terrain	1,000m × 1,000m
Total duration of sensing periods	10,000 s

5.4.1 Energy-Saving Using Compression

The energy expenditure in the transfer of data from the source node to the relay node can be calculated using equation 5.26.

Residual Energy=Initial Energy– Node's current energy (5.26)

In figure 5.7 the single-hop energy consumption between the source node and its relay is shown. Here five different source nodes are used. In figure 5.7 the distance between source node and relay is considered. As the distance increases the energy consumed also increases. POR consumes less energy with limited number of transmission whereas PRMTI consumes more energy with increased transmissions. Numerical values obtained by simulation are shown in table 5.4.

Table 5.4: Residual Energy – POR and PRMTI

Initiator node	Residual Energy - POR	Residual Energy - PRMTI
1	1.80 J	1.71 J
2	1.74 J	1.54 J
3	1.58 J	1.35 J
4	1.42 J	1.21 J
5	1.31 J	1.10 J

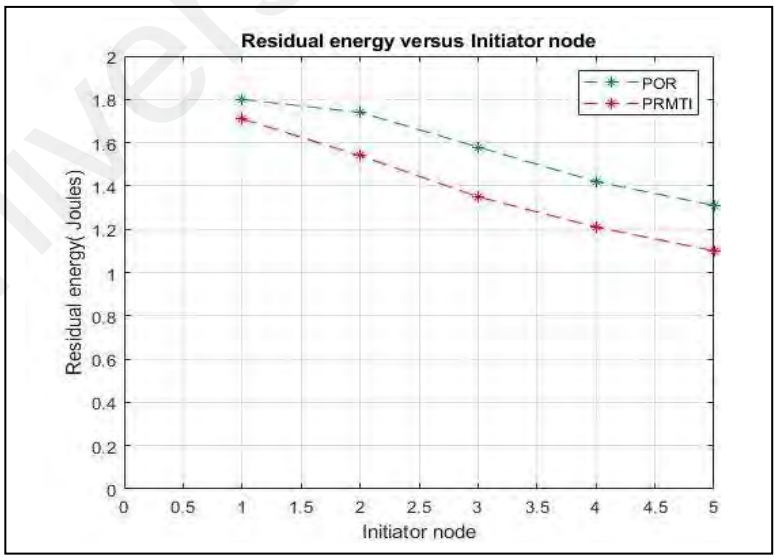

Figure 5.7: POR and PRMTI - Residual Energy

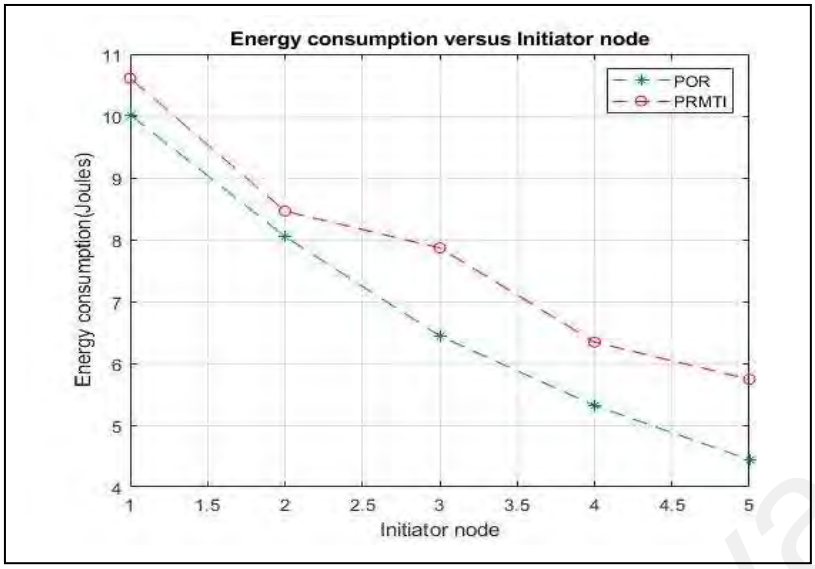

Figure 5.8: POR and PRMTI - Energy Consumption versus Initiator node

Table 5.5: Energy Consumption – POR and PRMTI

Initiator node	Energy Consumption POR	Energy Consumption PRMTI
1	10.018 J	10.61 J
2	8.05 J	8.461 J
3	6.441 J	7.866 J
4	5.316 J	6.341 J
5	4.441 J	5.741 J

Energy consumption associated with number of hops and forwarding nodes is calculated from source node. PRMTI due to excessive forwarding consumes more energy than POR is shown in figure 5.8. The energy expenditure from source node to relay node and its subsequent forwarding of data to the sink is calculated using equation 5.27. The numerical values of energy consumption are shown in table 5.5. The development of POR focuses on the transmission of data when the available wireless network resources are sparse.

$$\text{Energy consumption} = \frac{\text{Total energy consumed} \times \text{Number of hops}}{\text{Number of forwarding sensor nodes}} \qquad (5.27)$$

Hence sensor nodes forward data only when the activation function matches the target response. Thus the number of transmissions is reduced and energy expenditure is lower in POR. In PRMTI when the relay node exceeds the forwarding threshold then it transfers data. Hence the number of transmissions is increased and there is an increase in energy expenditure. In figure 5.9 the number of transmissions in PRMTI is higher considering the data forwarding capability. In POR number of transmissions is decreased and forwarding happens only if the target response is satisfied. The cost incurred for data transmission is calculated using table 5.6.

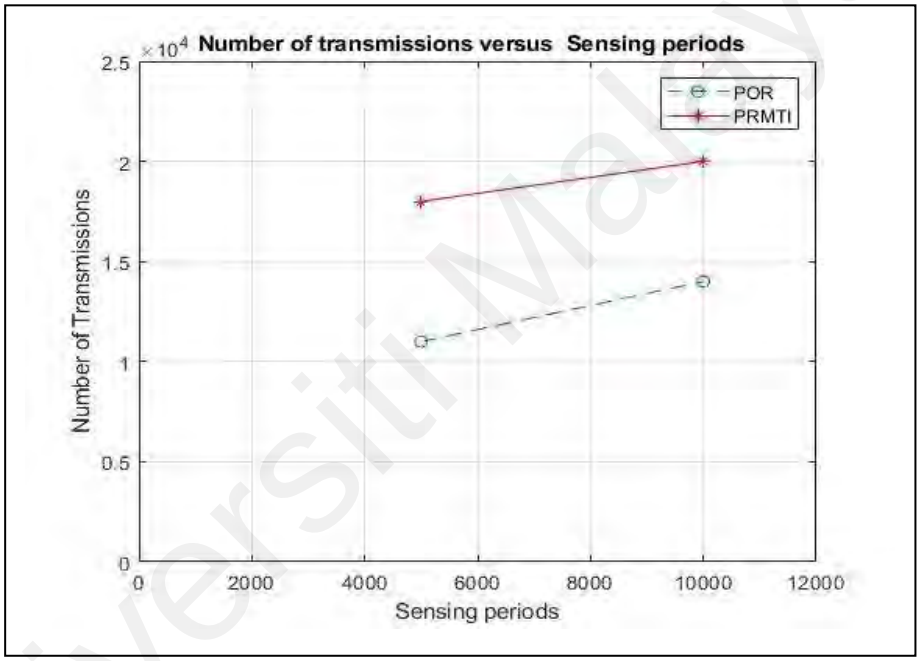

Figure 5.9: POR and PRMTI – Sensing period versus Number of Transmissions

Table 5.6: Sensing Period Vs Number of Transmissions - POR and PRMTI

Protocol	Sensing Period (s)	Number of Transmissions
POR	5,000	109,000
PRMTI	5,000	183,000
POR	10,000	129,000
PRMTI	10,000	197,000

5.4.2 Interpretation of Energy Consumption

If the number of transmissions increases there is a sudden depreciation in energy consumption due to continuous forwarding. Hence to denote the energy consumption two parameters namely residual energy and Energy consumption associated with number of hops is considered. In figure 5.10 a similar compression ratio is used in both the proposed protocols with two different forwarding criteria. The transmission cost in table 5.7 shows the recovery at the sink with the consolidated number of transmissions from the sensors.

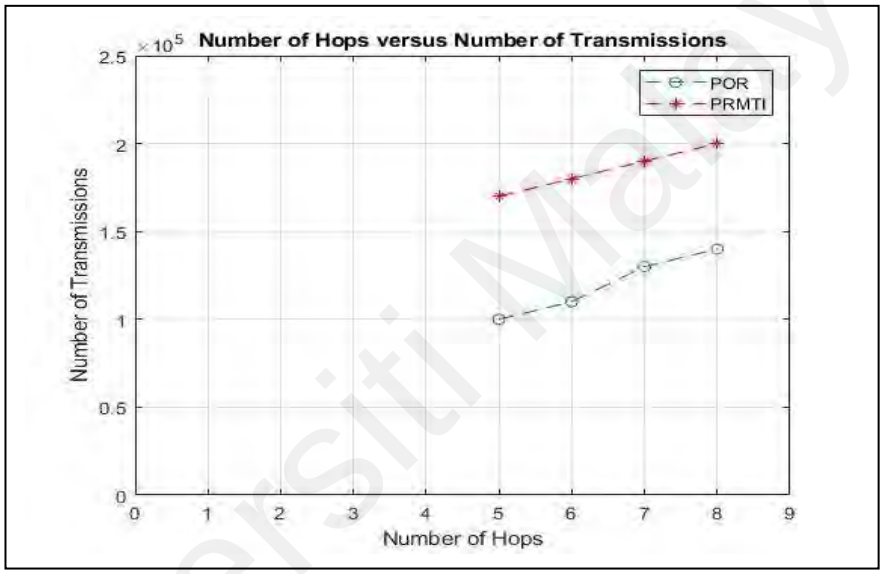

Figure 5.10: POR and PRMTI - Number of Hops versus Number of Transmissions

Table 5.7: Number of Hops Vs Number of Transmissions - POR and PRMTI

Number of Hops	No of Transmissions - POR	No of Transmissions - PRMTI
5	98,145	176,943
6	107,431	188,926
7	118,974	190,009
8	129,000	197,000

Sensors have less computational capability in forwarding data. The amount of information generated at the source sensor decreases as it approaches sink. Hence the number-of-hops metric is shown in figure 5.10 to relate the forwarding capabilities to minimal intermittent connectivity and communication void.

5.4.3 Comparisons of Proposed Protocols with Existing Protocols

Performance evaluation of residual energy and number of transmissions of proposed protocol (POR & PRMTI) with CDG (Luo et al., 2009) using the simulation parameters as in table 5.3 is shown in this section. CDG independently does compression followed by routing considering the global traffic with appropriate load in transmission and thereby persevering energy efficiency of certain nodes. In figure 5.11 five different nodes with different distance to sink are considered and the residual energy is calculated. The Residual energy of POR is better than CDG after several transmissions. Numerical value of residual energy of both protocols is shown in table 5.8.

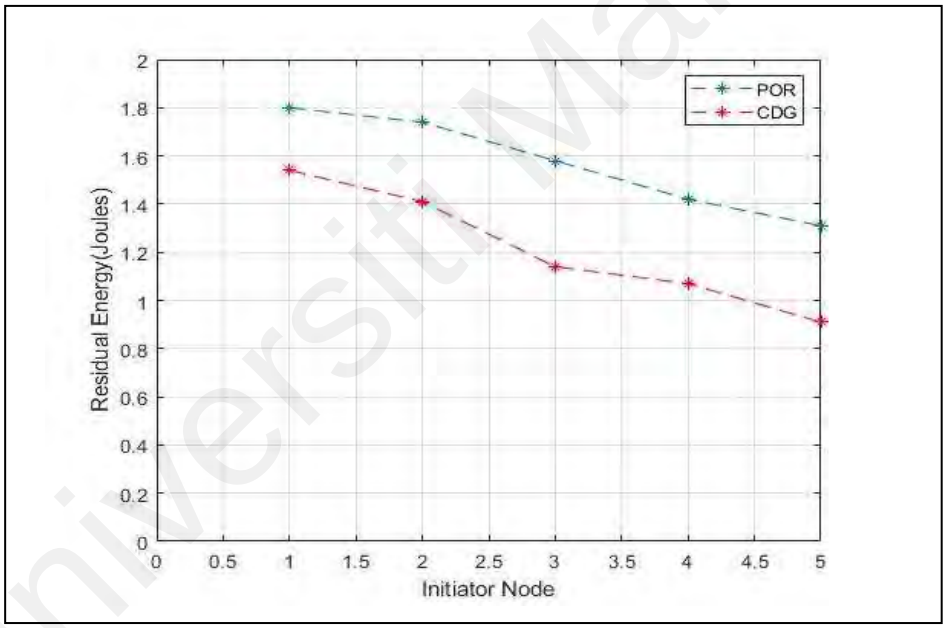

Figure 5.11: Comparison graph POR and CDG - Residual Energy

Table 5.8: Residual Energy - POR and CDG

Initiator node	Residual Energy - POR	Residual Energy - CDG
1	1.80 J	1.54 J
2	1.74 J	1.41 J
3	1.58 J	1.14 J
4	1.42 J	1.07 J
5	1.31 J	0.91 J

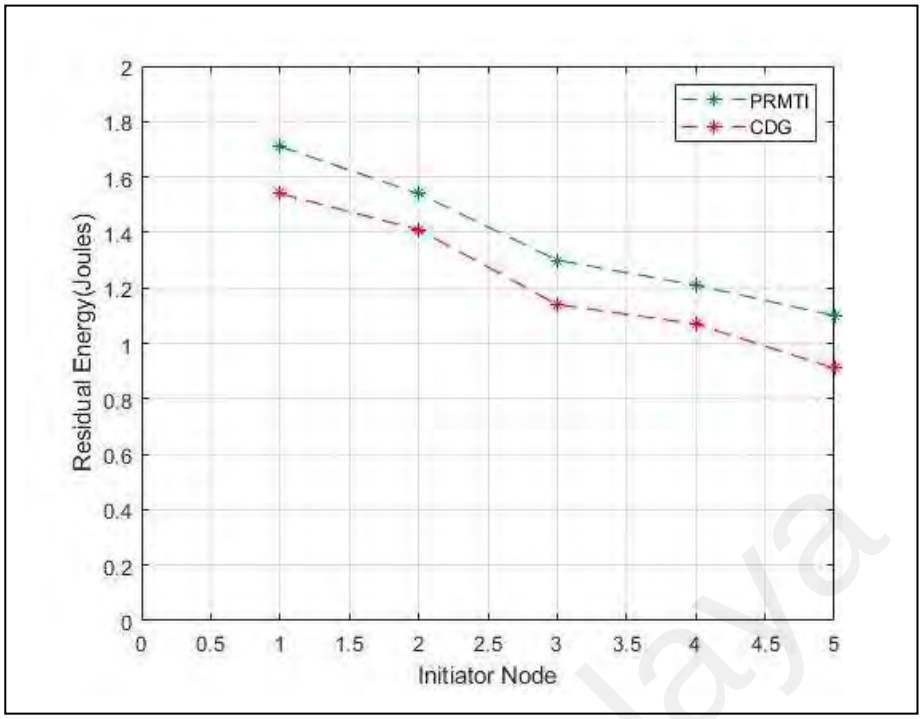

Figure 5.12: Comparison graph PRMTI and CDG - Residual Energy

Table 5.9: Residual Energy - PRMTI and CDG

Initiator node	Residual Energy - PRMTI	Residual Energy - CDG
1	1.71 J	1.54 J
2	1.54 J	1.41 J
3	1.35 J	1.14 J
4	1.21 J	1.07 J
5	1.10 J	0.91 J

In figure 5.12 five different nodes with different distance to sink are considered and their residual energy is calculated. The residual energy of PRMTI has been better even with increased number of transmission than CDG after several transmissions. Numerical value of residual energy of PRMTI and CDG protocol is shown in table 5.9. In figure 5.13 the number of transmissions versus sensing period is shown. The number of transmissions is considerably decreased in POR when compared to CDG. In the case of POR this is achieved by using intermediate node's processing capability balancing the global traffic. The values obtained with MATLAB simulations are shown in table 5.10.

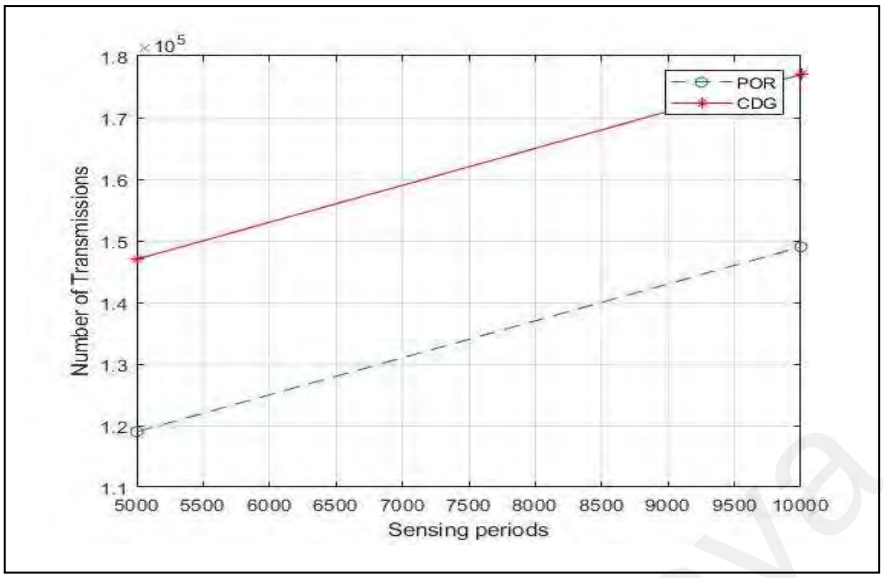

Figure 5.13: POR and CDG - Number of Transmissions

Table 5.10: Number of Transmissions - POR and CDG

Sensing period	POR Protocol Number of Transmissions	CDG Protocol Number of Transmissions
5,000	1,19,000	1,47,000
10,000	1,49,000	1,77,000

5.5 Data analysis of POR and PRMTI

In the previous works, the impact of transmissions with spatial and time coordinates and its results are inferred in Multinomial Logistic Regression analysis. Further results of error rate and energy consumption are statistically analyzed using SPSS (Zuhairy & Al., 2018). There are several works explaining the tabulated statistical results. They infer the impact of mobility and connectivity through Analysis of Variance (ANOVA) using SPSS (Ahmad et.al., 2018). Analogous to these, in our work the number of transmissions is analyzed using time series and residual energy consumption using Generalized Linear Model (GLM).

There are numerous algorithms used for statistical signal processing for different applications. However in this work data obtained via simulations are interpreted using time series analysis and by expert modeler method. Time series method is used to prove the increase in number of transmissions as the sensing period is increased. Expert modeler method proves the relationship between residual energy and overall energy consumption by interpreting the simulation results.

5.5.1 POR – Analysis on Number of Transmissions

The efficiency of POR and PRMTI is validated using a time series algorithm. The significance of time series analysis is it simultaneously associates the dependent and independent variables.

Null Hypothesis: There is no relationship between the sensing period and reduction in number of transmissions.

Alternate Hypothesis: There is a relationship between the sensing period and reduction in number of transmissions.

Table 5.11: Comparison of proposed protocols with Sensing Period

Sensing period (seconds)	POR	PRMTI
4000	98,145	176,943
6000	107,431	188,926
8000	118,974	190,009
10000	129,000	197,000

Table 5.12: Model description for numerical evaluations - POR.

Model Description		
Model id	Number of transmissions	Model Type ARIMA (0,0,0)

Table 5.13: Model fit statistics for Number of Transmissions – POR

Number of Transmissions Model	Stationary R squared	R squared	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC
	0.998	0.998	651.57	0.351	377.400	0.699	751.100	13.652

Table 5.14: Residual ACF summary - POR

Lag	Mean
Lag 1	-0.669
Lag 2	0.171
Lag 3	-0.002

Table 5.15: Residual PACF summary - POR

Lag	Mean
Lag 1	-0.669
Lag 2	0.171
Lag 3	-0.002

Table 5.11 shows the comparison of POR and PRMTI at different sensing periods. Table 5.12 denotes the appropriate model for Auto Regressive Integrated Moving Average. Table 5.13 shows the following model fit statistics. The optimal order of differencing can be found with Root Mean Square Error (RMSE). Diagnostic checking can be done with Bayesian Information criteria (BIC), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Maximum Absolute Percentage Error (MaxAPE), Maximum Absolute Error (MaxAE). Residual value with ACF and PACF summary states the model fits is appropriate with white noise.

The Auto Correlation Function (ACF) denotes the simple correlation between the current observation and the observation is “p” period from current observation. It is shown in table 5.14. The Partial Auto Correlation Function (PACF) denotes the degree of association between the current observation and the observation is “p” period from current observation. It is shown in table 5.15. The figure 5.14 denotes the model fit statistics values of stationary R square versus frequency for POR. The figure 5.15 shows the model fit statistics values of R squared versus frequency for POR.

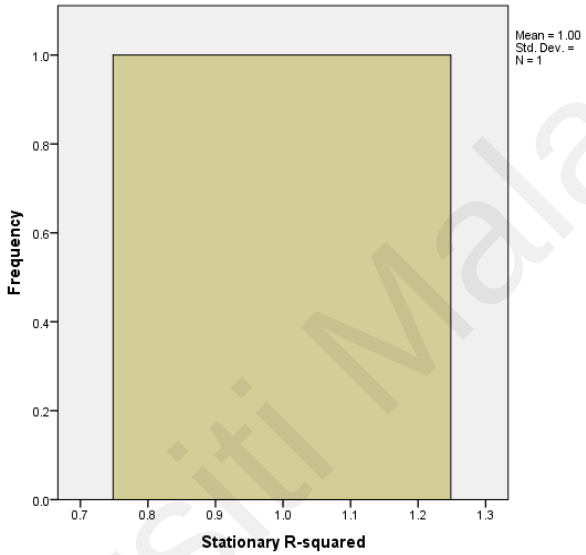

Figure 5.14: POR - Stationary R squared plot

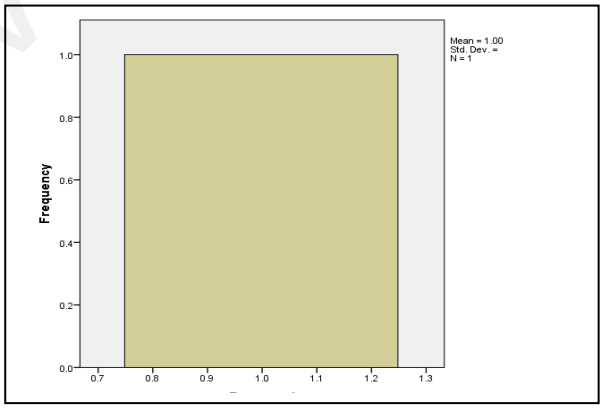

Figure 5.15: POR - R squared plot

The figure 5.16 denotes the model fit statistics values of RMSE versus frequency for POR. The figure 5.17 denotes the model fit statistics values of Normalized BIC versus frequency for POR. Table 5.16 shows the predicted, lower and upper limits of data interpreted using simulation results of data analysis. This is just an association of transmission cost with sensing period obtained with Auto Regressive Integrated Moving Average (ARIMA) model for POR.

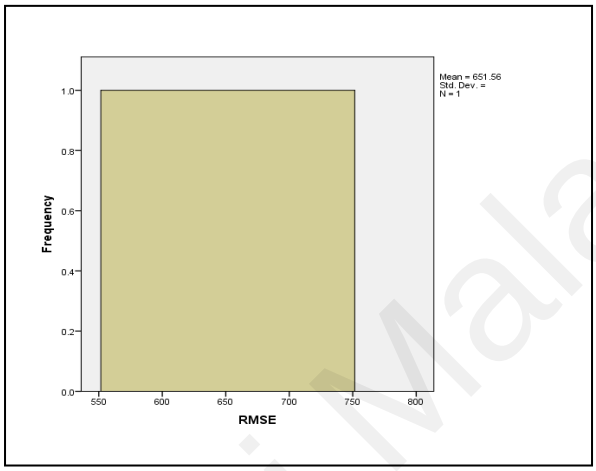

Figure 5.16: POR - RMSE versus Frequency

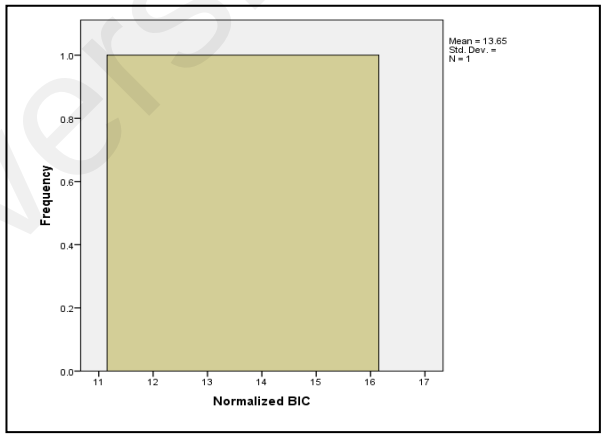

Figure 5.17: POR - Normalized BIC versus Frequency

Table 5.16: Predicted, Lower and Upper Confident Limit – POR

Predicted	Lower Confident Limit	Upper Confident Limit
97771	94968	100575
108182	105379	110986
118593	115789	121396
129004	126200	131807

5.5.2 PRMTI – Analysis on Number of Transmissions

Table 5.17 denotes the appropriate model of Auto Regressive Integrated Moving Average for PRMTI. Table 5.18 shows the model fit statistics values. The Auto Correlation Function (ACF) denotes the simple correlation between the current observation and the observation is "p" period from current observation. It is shown in table 5.19 for PRMTI. The Partial Auto Correlation Function (PACF) denotes the degree of association between the current observation and the observation is "p" period from current observation. It is shown in table 5.20.

Table 5.17: Model description for numerical evaluations - PRMTI

Model Description		
Model id	Number of transmissions	Model Type ARIMA (0,0,0)

Table 5.18: Model fit statistics for Number of Transmissions - PRMTI

Number of Transmissions Model	Stationary R squared	R squared	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC
	0.902	0.902	3190.255	1.013	1884.660	1.995	3769.200	16.829

Table 5.19: Residual ACF summary - PRMTI

Lag	Mean
Lag 1	- 0.597
Lag 2	0.55
Lag 3	0.42

Table 5.20: Residual PACF summary - PRMTI

Lag	Mean
Lag 1	- 0.597
Lag 2	0.468
Lag 3	-0.376

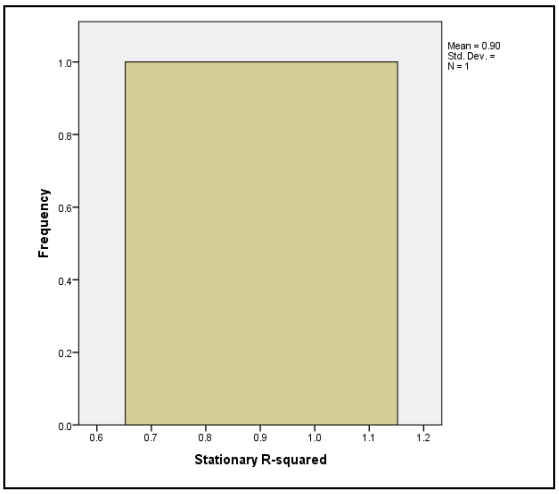

Figure 5.18: PRMTI – Stationary R-squared

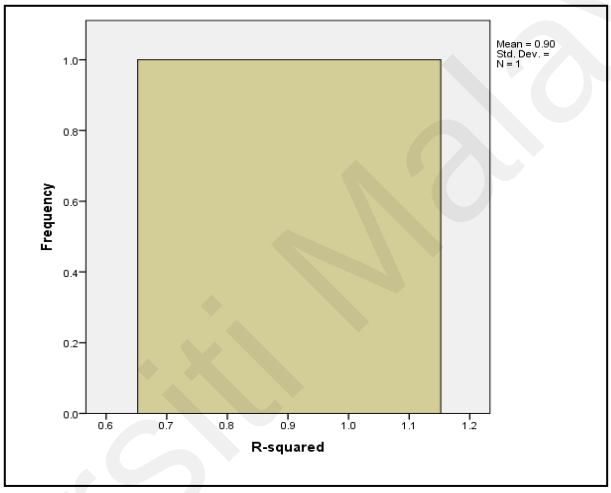

Figure 5.19: PRMTI- R squared plot

The figure 5.18 denotes the model fit statistics values of stationary R square versus frequency for PRMTI protocol. The figure 5.19 denotes the model fit statistics values of R square versus frequency for PRMTI protocol. The figure 5.20 denotes the model fit statistics values of RMSE square versus frequency for PRMTI. Figure 5.21 shows the normalized BIC values for PRMTI. The table 5.21 shows the predicted, lower and upper limits of data interpreted using simulation results. This is just an association of transmission cost with sensing period obtained using Auto Regressive Integrated Moving Average (ARIMA) model for PRMTI

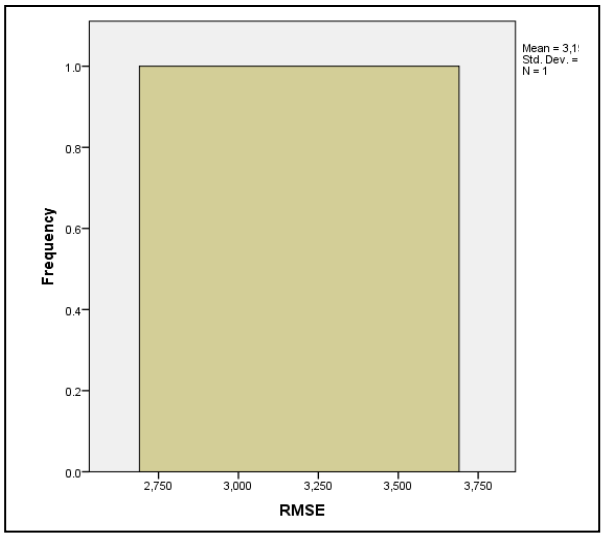

Figure 5.20: PRMTI - RMSE versus Frequency

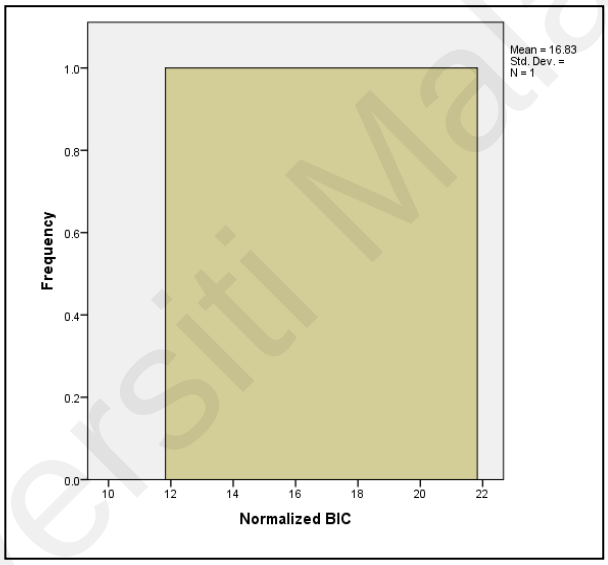

Figure 5.21: PRMTI - Normalized BIC versus Frequency

Table 5.21: Predicted, Lower and Upper Confident Limit - PRMTI

Predicted	Lower Confident Limit	Upper Confident Limit
179031	165305	192758
185157	171430	198883
191282	177556	205009
197408	183681	211134

From the numerical evaluations it is clearly evident, POR yields an association indicating the alternate hypothesis is proved. Simulation results of both the protocol indicate the decrease in number of transmissions with increase in sensing period. Since sensors are subjected to varying observation cost, determining the transmission cost estimation is done in three intervals namely predicted, lower and upper confident limits. The development of POR in a resource constrained environment in table 5.16 shows lower values when compared to table 5.21 (PRMTI) developed in a resource unconstrained environment.

5.5.3 POR – Analysis on Energy Consumption

The analysis on energy consumption is performed using Expert Modeler. There are two interpretation of energy consumption the first is residual energy of individual node and other is the overall energy incurred for transfer of data to sink associated with number of hops. The death of an individual sensor node can partition the network into two. The values in the above table represents the individual energy of sensor nodes denoted as residual energy (RE). The overall energy consumption occurred in transmission with respect to the number of hops is denoted as EC. So the overall energy consumption is a factor depending on several nodes within the sensing field. The residual energy and overall energy consumption of POR and RMTI is shown in table 5.22.

Table 5.22: Energy Consumption analysis using Expert modeler

Initiator Node	POR_RE	POR_EC	PRMTI_RE	PRMTI_EC
1	1.80 J	10.018	1.71	10.61
2	1.74 J	8.05	1.54	8.461
3	1.58 J	6.441	1.35	7.866
4	1.42 J	5.316	1.21	6.341
5	1.31 J	4.441	1.10	5.741

The variables used in expert modeler are represented as follows **POR_RE:** Residual Energy of Perceptron-based Optimal Routing, **POR_EC:** Energy consumption of Perceptron-based Optimal Routing, **PRMTI_RE:** Residual Energy of Perceptron-based Routing with Moderate Traffic Intensity, **PRMTI_EC:** Energy consumption of Perceptron-based Routing with Moderate Traffic Intensity.

Null Hypothesis: There is no relationship between the residual energy of individual node and overall energy consumption.

Alternate Hypothesis: There is a relationship between the residual energy and overall energy consumption.

The simulations in this work are carried out using expert modeler where only dependant variables are used. In this case the dependant variables used are residual energy and energy consumption of the proposed protocols. The table 5.23 indicates the model description of residual energy of sensor nodes a Holt method is to be performed on the described model. Table 5.24 shows the summary of POR_RE with its model fit statistics. The Table 5.25 shows the summary of POR_RE with exponential smoothing model and it is the best fitting model.

Table 5.23: Model description for numerical evaluations - POR_RE

Model Description	
Model id POR_RE Model_1	Model Type used Holt

Table 5.24: Model summary - POR_RE

Number of Transmissions Model	R squared	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC
	0.973	0.039	1.436	0.24	2.686	0.47	-5.834

Table 5.25: Exponential smoothing model parameters - POR_RE

	Estimate	SE	t	Sig
Alpha (Level)	**0.392**	**0.254**	**1.545**	**0.220**
Gamma (Trend)	**7.194E-6**	0.105	6.860E-5	1.000

Table 5.26: Predicted, Lower and Upper Confident Limit - POR_RE.

Predicted	Lower Confident Limit	Upper Confident Limit
1.838	1.713	1.963
1.693	1.568	1.818
1.582	1.457	1.706
1.451	1.326	1.576
1.309	1.184	1.434

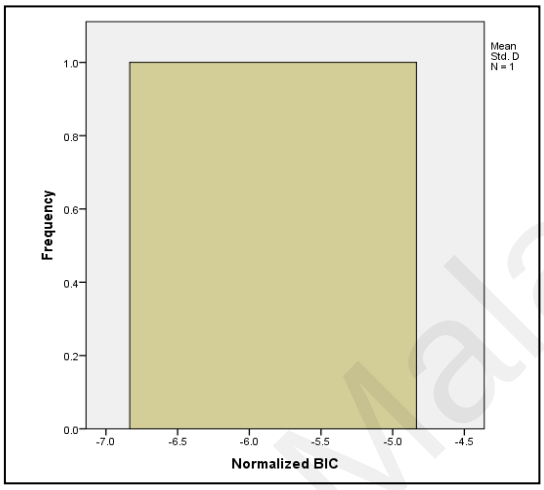

Figure 5.22: POR_RE - Normalized BIC versus Frequency

Table 5.27: Model description for numerical evaluations - POR_EC

Model Description	
Model id POR_EC Model_1	Model Type used Holt

Table 5.28: Model summary - POR_EC

Number of Transmissions Model	R squared	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC
	0.978	0.377	4.063	0.219	9.108	0.484	-1.308

The smaller value of normalized BIC in the exponential smoothing table indicates the POR_RE values fitting the data considerably. Thus there is minimal penalty added with POR_RE. The table 5.26 represents the predicted, lower and upper limits of data interpreted using the simulation results. Figure 5.22 denotes the model fit statistics values of Normalized BIC versus frequency for POR_RE. Table 5.27 indicates the model description of Holt model. The table 5.28 shows the model summary for POR_RE with its model fit statistics.

Table 5.29 shows the model summary of POR_EC with exponential smoothing model and it is the best fitting model. The smaller value of normalized BIC in the exponential smoothing table indicates POR_EC values fits the data considerably. Thus there is minimal penalty added with POR_EC. The table 5.30 shows the energy consumption with upper and lower confidence limits with the values of POR_EC using expert modeler. The figure 5.23 denotes the model fit statistics values of normalized BIC versus frequency of POR_EC.

Table 5.29: Exponential smoothing model parameters - POR_EC

	Estimate	SE	T	Sig
Alpha (Level)	1.000	.759	1.318	.279
Gamma (Trend)	1.000	1.119	.893	.437

Table 5.30: Predicted, Lower and Upper Confident Limit - POR_EC

Predicted	Lower Confident Limit	Upper Confident Limit
10.02	8.82	11.22
8.05	6.85	9.25
6.08	4.88	7.28
4.83	3.63	6.03
4.19	2.99	5.39

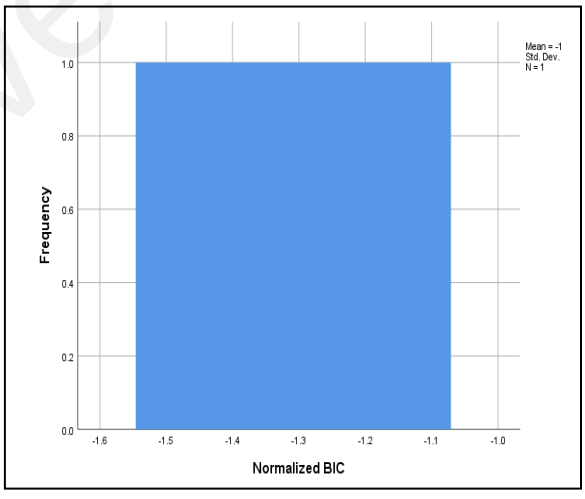

Figure 5.23: POR-EC - Normalized BIC versus Frequency

5.5.4 PRMTI- Analysis on Energy Consumption

Table 5.31 gives the model description of PRMTI_RE. Table 5.32 indicates the data of residual energy of sensor nodes performed with Expert modeler. Table 5.32 shows the model summary of PRMTI_RE with its model fit statistics. Table 5.33 shows the summary of PRMTI_RE with exponential smoothing model. Table 5.34 shows the predicted, lower and upper limits of data interpreted for PRMTI_RE. This is just an association of transmission cost with sensing period obtained with Auto Regressive Integrated Moving Average (ARIMA) model for PRMTI.

Table 5.31: Model description for numerical evaluations - PRMTI_RE

Model Description	
Model id PRMTI_RE Model_1	Model Type used Holt

Table 5.32: Model summary - PRMTI_RE

Stationary R squared	R squared	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC
-0.338	0.985	0.035	1.695	0.021	3.818	0.042	-6.066

Table 5.33: Exponential smoothing model parameters - PRMTI_RE

	Estimate	SE	t	Sig
Alpha (Level)	0.490	0.596	0.822	0.471
Gamma (Trend)	3.283E-7	0.092	3.563E-6	1.000

Table 5.34: Predicted, Lower and Upper Confident Limit - PRMTI_RE

Predicted	Lower Confident Limit	Upper Confident Limit
1.704	1.592	1.815
1.552	1.441	1.663
1.391	1.280	1.502
1.216	1.105	1.327
1.058	0.947	1.169

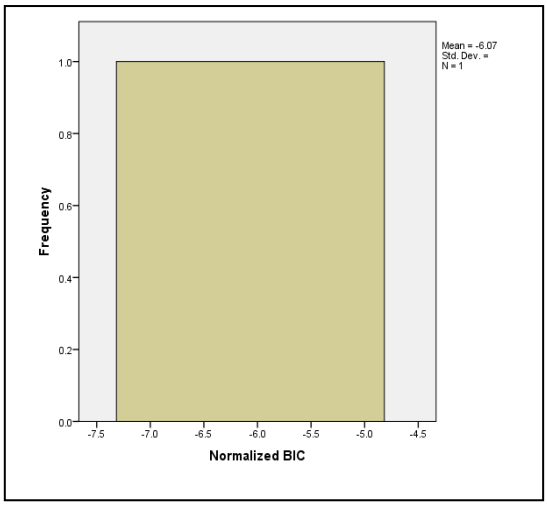

Figure 5.24: PRMTI_RE - Normalized BIC versus Frequency

Table 5.35: Model description for numerical evaluations - PRMTI_EC

Model Description	
Model id PRMTI_EC Model_1	Model Type used Holt

Table 5.36: Model summary - PRMTI_EC

Stationary R squared	R squared	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC
0.922	0.934	0.570	4.986	0.373	9.480	.802	-0.480

Table 5.37: Exponential smoothing model parameters - PRMTI_EC

	Estimate	SE	t	Sig
Alpha (Level)	0.392	0.470	0.835	0.465
Gamma (Trend)	3.052E-6	0.172	2.03E-5	1.000

Figure 5.24 denotes the model fit statistics values of Normalized BIC versus frequency for PRMTI_RE. The table 5.35 indicates the model for energy consumption of sensor nodes. Table 5.36 shows the model summary of PRMTI_EC with its model fit statistics. The Table 5.37 shows the exponential smoothing model and is proved as the best fitting model for PRMTI_EC.

Table 5.38: Predicted, Lower and Upper Confident Limit - PRMTI_EC

Predicted	Lower Confident Limit	Upper Confident Limit
10.345	8.530	12.159
9.263	7.449	11.078
7.763	5.948	9.577
6.617	4.803	8.432
5.323	3.509	7.138

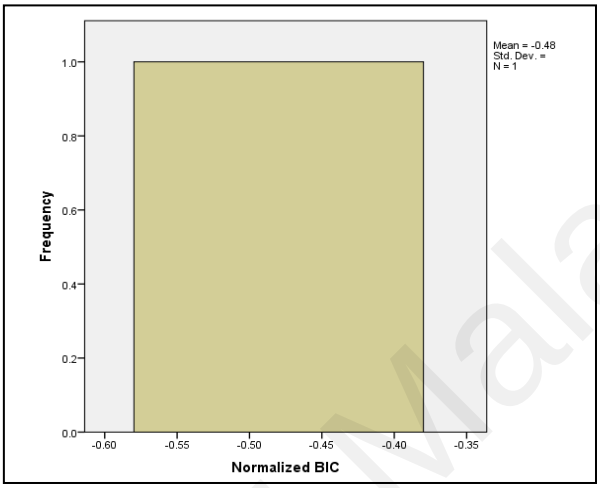

Figure 5.25: PRMTI_EC - Normalized BIC versus Frequency

The table 5.38 shows the predicted, lower and upper limits of data interpreted using simulation results of PRMTI_EC. This is just an association of transmission cost with sensing period obtained using Auto Regressive Integrated Moving Average (ARIMA) model for PRMTI. Figure 5.25 depicts the PRMTI_EC - Normalized BIC versus Frequency. It denotes the model fit statistics values of Normalized BIC versus frequency of PRMTI_EC.

The smaller value of normalized BIC in the exponential smoothing table indicates PRMTI_RE values fits the data considerably. However while comparing both PRMTI_RE with lower BIC. There is a strong relationship between the residual energy and overall energy consumption. Thus the results indicate alternate hypothesis is proved.

5.6 Conclusion

The efficiency of compressive sensing in wireless sensor nodes not only depends on sensed data but also on the minimum number of transmissions needed to reach the sink. The framework has two protocols with different forwarding strategies in the transfer of data to the sink. Both protocols incorporate discrete wavelet transform in order to ensure the sparsity of sensed data and recovery at the sink. The framework of Perceptron-based Optimal Routing (POR) and Perceptron-based Routing with Moderate Traffic Intensity (PRMTI) initiates data forwarding and do not require a long training phase. Data transmission subsequently takes place by considering the available bandwidth and the concurrent transmission capability of the intermediate node involved in forwarding. Hence load balancing is achieved in both the algorithms. POR uses a bipolar logic with a target value at the relay node for forwarding. If the target response is not satisfactory then weight update is performed iteratively until the incoming data meets the target response. PRMTI forwards when the link capacity exceeds the defined data- forwarding threshold at the relay node. Thus the transfer of compressed data in unreliable wireless links is performed with binary classifiers using the threshold values resulting in a fully connected topology. The biasing effect of the framework with its associated metrics for data forwarding shows greater transmission efficiency for compressed data in a wireless environment. The simulation outcome with regard to residual energy, energy consumption and the number of transmissions is estimated for the proposed framework. The comparison of the proposed framework has been performed with an existing protocol CDG. The frameworks show increased residual energy and reduction in number of transmissions than CDG. Time series analysis has been carried out on the simulation results of proposed framework in order to ascertain the reduction in the number of transmissions within the allocated sensing period. The simulation results on energy consumption is validated with expert modeler determining the relationship between residual energy of individual nodes and overall energy consumption.

CHAPTER 6: CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

The significance of the thesis is to overcome the constraints involved in sensing and communication cost of wireless sensor nodes. The first framework embodies solutions for energy efficiency to prolong the lifetime of sensors. Cluster members can be elected as cluster head based on the main indices favoring energy. The multi-hop communication is performed in varying sensing field and the packet exchange occurs either to static or mobile sink based on the nearest distance. The mobile sink preserves both, one-hop neighbors of static sink and peripheral nodes within the sensing field.

The penultimate framework incorporates compressive sensing. It reduces the sensing cost by sparse sensed data representation at source and post processing the sparse data at sink. Data forwarding from source to neighbor using spatial-temporal coordinates and correlation coefficients ensures the minimal transmission and reconstruction accuracy being achieved. The sensing cost is reduced by compressive sensing and the number of transmissions is reduced during data forwarding through correlation coefficients ensuring energy proficiency.

Final framework ensures reduction in transmission cost and sensing cost. Sensing cost is reduced by incorporating compressive sensing however wireless link capacity may degrade due to traffic intensity wherein intermediate nodes are unable to forward data to sink. Hence channel impairment between source node and intermediate nodes are interpreted for forwarding data. To analyze the link capacity and traffic intensity single layer perceptron-based framework for data forwarding has been incorporated. The framework constitutes two solutions. In the first solution when there is a match in data forwarding threshold value between source and intermediate forwarder then forwarding happens. This solution suits for resource constrained networks.

In the second solution when any of the forwarder exceeds the minimal data forwarding threshold value, forwarding happens from source node to intermediate forwarder and it is applicable for resourceful networks.

6.2 Future Work

Future work will focus on development of algorithm for sensing layer and for those nature of applications where wireless sensor nodes is deployed. The sensing layer provides better techniques to achieve sparsity and minimal power consumption in data acquisition process. The sensing layer in line with communication operation transmits the data from sensors, for the post processing process and reconstruction process at sink within the estimated time period. Communication operation would be interfaced to IOT platforms. The nodal activity and channel response for signal elevation or the de-elevation strategy would be matched in transmission. Development of alternate path-based communication prior to sensing can be inferred depending on the application of sensor's sensitive or insensitive to delay. Further direction of research can be obtained by exploring the spatial sparsity liable on the application involved in deployment phase. This spatial sparsity has to be associated with channel state information ensuring high reconstruction accuracy.

REFERENCES

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: a survey. Computer networks, 38(4), 393-422.

Aeron, S., Zhao, M., & Saligrama, V. (2006). Fundamental trade-offs between sparsity, sensing diversity and sensing capacity. In 2006 Fortieth Asilomar Conference on Signals, Systems and Computers (pp. 295-299). IEEE.

Anastasi, G., Conti, M., Di Francesco, M., & Passarella, A. (2009). Energy conservation in wireless sensor networks: A survey. Ad hoc networks, 7(3), 537-568.

Abbasi, A. A., & Younis, M. (2007). A survey on clustering algorithms for wireless sensor networks. Computer communications, 30(14-15), 2826-2841.

Arjoune, Y., Kaabouch, N., El Ghazi, H., & Tamtaoui, A. (2018). A performance comparison of measurement matrices in compressive sensing. International Journal of Communication Systems, 31(10).

Aeron, S., Zhao, M., & Saligrama, V. (2008). Algorithms and bounds for sensing capacity and compressed sensing with applications to learning graphical models. In 2008 Information Theory and Applications Workshop (pp. 303-309). IEEE.

Abuarqoub, A., Hammoudeh, M., Adebisi, B., Jabbar, S., Bounceur, A., & Al-Bashar, H. (2017). Dynamic clustering and management of mobile wireless sensor networks. Computer Networks, 117, 62-75.

Amjad, M., Afzal, M. K., Umer, T., & Kim, B. S. (2017). QoS-aware and heterogeneously clustered routing protocol for wireless sensor networks. IEEE Access, 5, 10250-10262.

Abo-Zahhad, M. M., Hussein, A. I., & Mohamed, A. M. (2015). Compressive sensing algorithms for signal processing applications: A survey. International Journal of Communications, Network and System Sciences, 8(06), 197.

Acimovic, J., Beferull-Lozano, B., &Cristescu, R. (2005). Adaptive distributed algorithms for power-efficient data gathering in sensor networks. In 2005 International Conference on Wireless Networks, Communications and Mobile Computing (Vol. 2, pp. 946-951). IEEE.

Alsheikh, M. A., Lin, S., Niyato, D., & Tan, H. P. (2014). Machine learning in wireless sensor networks: Algorithms, strategies, and applications. IEEE Communications Surveys & Tutorials, 16(4), 1996-2018.

Alsheikh, M. A., Hoang, D. T., Niyato, D., Tan, H. P., & Lin, S. (2015). Markov decision processes with applications in wireless sensor networks: A survey. IEEE Communications Surveys & Tutorials, 17(3), 1239-1267.

Alsheikh, M. A., Lin, S., Niyato, D., & Tan, H. P. (2016). Rate-distortion balanced data compression for wireless sensor networks. IEEE Sensors Journal, 16(12), 5072-5083.

Amarlingam, M., Mishra, P. K., Rajalakshmi, P., Channappayya, S. S., & Sastry, C. S. (2018). Novel Light Weight Compressed Data Aggregation using sparse measurements for IoT networks. Journal of Network and Computer Applications, 121, 119-134.

Alwan, N. A., & Hussain, Z. M. (2019). Compressive sensing with chaotic sequences: an application to localization in wireless sensor networks. Wireless Personal Communications, 105(3), 941-950.

Anand, J. V., & Titus, S. (2017). Energy efficiency analysis of effective hydrocast for underwater communication. International Journal of Acoustics & Vibration, 22(1) 44-50.

Ahmad, M., Li, T., Khan, Z., Khurshid, F., & Ahmad, M. (2018). A Novel Connectivity-Based LEACH-MEEC Routing Protocol for Mobile Wireless Sensor Network. Sensors, 18(12), 4278.

Bult, K., Burstein, A., Chang, D., Dong, M., Fielding, M., Kruglick, E., & Marcy, H. (1996). Low power systems for wireless microsensors. In Proceedings of 1996 International Symposium on Low Power Electronics and Design (pp. 17-21). IEEE.

Bader, S., Ma, X., & Oelmann, B. (2014). On the modeling of solar-powered wireless sensor nodes. Journal of Sensor and Actuator Networks, 3(3), 207-223.

Bhardwaj, M., Garnett, T., & Chandrakasan, A. P. (2001). Upper bounds on the lifetime of sensor networks. In ICC 2001. IEEE International Conference on Communications. Conference Record (Cat. No. 01CH37240) (Vol. 3, pp. 785-790). IEEE.

Baraniuk, R. G., Cevher, V., Duarte, M. F., &Hegde, C. (2008). Model-based compressive sensing. IEEE Transactions on Information Theory 56, (4) 1982 – 2001.

Bawane, P., & Kannu, A. P. (2019). Time varying sparse support recovery. Signal Processing, 161, 214-226.

Berger, C. R., Wang, Z., Huang, J., & Zhou, S. (2010). Application of compressive sensing to sparse channel estimation. IEEE Communications Magazine, 48(11), 164- 174.

Baradaran, A. A., & Navi, K. (2017). CAST-WSN: The presentation of new clustering algorithm based on Steiner tree and C-means algorithm improvement in wireless sensor networks. Wireless Personal Communications, 97(1), 1323-1344.

Bettstetter, C. (2002). On the minimum node degree and connectivity of a wireless multihop network. In Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing (pp. 80-91). ACM.

Blumensath, T., & Davies, M. E. (2009). Iterative hard thresholding for compressed sensing. Applied and computational harmonic analysis, 27(3), 265-274.

Baccour, N., Koubâa, A., Mottola, L., Zúñiga, M. A., Youssef, H., Boano, C. A., & Alves, M. (2012). Radio link quality estimation in wireless sensor networks: A survey. ACM Transactions on Sensor Networks (TOSN), 8(4), 34.

Bajwa, W., Haupt, J., Sayeed, A., & Nowak, R. (2006). Compressive wireless sensing. In Proceedings of the 5th international conference on Information processing in sensor networks (pp. 134-142). ACM.

Blanchard, J. D., Cartis, C., & Tanner, J. (2011). Compressed sensing: How sharp is the restricted isometry property? SIAM review, 53(1), 105-125.

Chen, B., Tong, L., & Varshney, P. K. (2006). Channel-aware distributed detection in wireless sensor networks. IEEE Signal Processing Magazine, 23(4), 16-26.

Cetin, M., Chen, L., Fisher, J. W., Ihler, A. T., Moses, R. L., Wainwright, M. J., & Willsky, A. S. (2006). Distributed fusion in sensor networks. IEEE Signal Processing Magazine, 23(4), 42-55.

Choi, J. W., Shim, B., Ding, Y., Rao, B., & Kim, D. I. (2017). Compressed sensing for wireless communications: Useful tips and tricks. IEEE Communications Surveys & Tutorials, 19(3), 1527-1550.

Chong, C. Y., & Kumar, S. P. (2003). Sensor networks: evolution, opportunities, and challenges. Proceedings of the IEEE, 91(8), 1247-1256.

Cardei, M., Wu, J., Lu, M., &Pervaiz, M. O. (2005). Maximum network lifetime in wireless sensor networks with adjustable sensing ranges. In WiMob'2005), IEEE International Conference on Wireless and Mobile Computing, Networking And Communications, 2005. (Vol. 3, pp. 438-445). IEEE.

Cheng, S., Cai, Z., & Li, J. (2014). Curve query processing in wireless sensor networks. IEEE Transactions on Vehicular Technology, 64(11), 5198-5209.

Cristescu, R., Beferull-Lozano, B., & Vetterli, M. (2005). Networked Slepian-Wolf: theory, algorithms, and scaling laws. IEEE Transactions on Information Theory, 51(12), 4057-4073.

Chen, D., &Varshney, P. K. (2004). QoS Support in Wireless Sensor Networks: A Survey. In International conference on wireless networks (Vol. 233, pp. 1-7).

Cao, Q., He, T., Fang, L., Abdelzaher, T., Stankovic, J., & Son, S. (2006). Efficiency centric communication model for wireless sensor networks. In Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications (pp. 1-12). IEEE.

Chen, Y., & Zhao, Q. (2005). On the lifetime of wireless sensor networks. IEEE Communications letters, 9(11), 976-978.

Cheng, Z., Perillo, M., & Heinzelman, W. B. (2008). General network lifetime and cost models for evaluating sensor network deployment strategies. IEEE Transactions on mobile computing, 7(4), 484-497.

Chen, Q., Kanhere, S. S., & Hassan, M. (2009). Analysis of per-node traffic load in multi-hop wireless sensor networks. IEEE transactions on wireless communications, 8(2), 958-967.

Candes, E. J., Romberg, J. K., & Tao, T. (2006). Stable signal recovery from incomplete and inaccurate measurements. Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences, 59(8), 1207-1223.

Candès, E. J., & Wakin, M. B. (2008). An introduction to compressive sampling a sensing/sampling paradigm that goes against the common knowledge in data acquisition]. IEEE signal processing magazine, 25(2), 21-30.

Candes, E. J., Wakin, M. B., & Boyd, S. P. (2008). Enhancing sparsity by reweighted ℓ 1 minimization. Journal of Fourier analysis and applications, 14(5-6), 877-905.

Candes, E. J., Eldar, Y. C., Needell, D., & Randall, P. (2011). Compressed sensing with coherent and redundant dictionaries. Applied and Computational Harmonic Analysis, 31(1), 59-73.

Candes, E., & Romberg, J. (2007). Sparsity and incoherence in compressive sampling. Inverse problems, 23(3), 969.

Chepuri, S. P., & Leus, G. (2016). Sparse sensing for statistical inference. Foundations and Trends in Signal Processing, 9, (3–4), 233-368.

Cervin, A., Henriksson, D., Lincoln, B., Eker, J., & Arzen, K. E. (2003). How does control timing affect performance? Analysis and simulation of timing using Jitterbug and True Time. IEEE control systems magazine, 23(3), 16-30.

Chakrabarti, A., Sabharwal, A., &Aazhang, B. (2003). Using predictable observer mobility for power efficient design of sensor networks. In Information Processing in Sensor Networks (pp. 129-145). Springer, Berlin, Heidelberg.

Chang, J. H., & Tassiulas, L. (2004). Maximum lifetime routing in wireless sensor networks. IEEE/ACM Transactions on networking, 12(4), 609-619.

Chen, X., Sobhy, E. A., Yu, Z., Hoyos, S., Silva-Martinez, J., Palermo, S., & Sadler, B. M. (2012). A sub-Nyquist rate compressive sensing data acquisition front-end. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2(3), 542-551.

Cheng, J., Ye, Q., Jiang, H., Wang, D., & Wang, C. (2013). STCDG: An efficient data gathering algorithm based on matrix completion for wireless sensor networks. IEEE Transactions on Wireless Communications, 12(2), 850-861.

Chen, Z., Ranieri, J., Zhang, R., & Vetterli, M. (2015). DASS: Distributed adaptive sparse sensing. IEEE Transactions on Wireless Communications, 14(5), 2571-2583.

Chen, W., & Wassell, I. J. (2015). A decentralized Bayesian algorithm for distributed compressive sensing in networked sensing systems. IEEE Transactions on Wireless Communications, 15(2), 1282-1292.

Chen, W., Rodrigues, M. R., & Wassell, I. J. (2012). A frechet mean approach for compressive sensing date acquisition and reconstruction in wireless sensor networks. IEEE Transactions on wireless communications, 11(10), 3598-3606.

Chen, W., & Wassell, I. J. (2012). Energy-efficient signal acquisition in wireless sensor networks: a compressive sensing framework. IET wireless sensor systems, 2(1), 1-8.

Ciancio, A., Pattem, S., Ortega, A., & Krishnamachari, B. (2006). Energy-efficient data representation and routing for wireless sensor networks based on a distributed wavelet compression algorithm. In Proceedings of the 5th international conference on Information processing in sensor networks (pp. 309-316). ACM.

Crovella, M., & Kolaczyk, E. (2003). Graph wavelets for spatial traffic analysis. In IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428) (Vol. 3, pp. 1848-1857). IEEE.

Chen, S., Liu, J., Wang, K., & Wu, M. (2019). A hierarchical adaptive spatio-temporal data compression scheme for wireless sensor networks. Wireless Networks, 25(1), 429-438.

Chen, S., Zhang, S., Zheng, X., & Ruan, X. (2019). Layered adaptive compression design for efficient data collection in industrial wireless sensor networks. Journal of Network and Computer Applications, 129, 37-45.

Deshpande, A., Guestrin, C., Madden, S. R., Hellerstein, J. M., & Hong, W. (2004). Model-driven data acquisition in sensor networks. In Proceedings of the Thirtieth international conference on Very large data bases-Volume 30(pp. 588-599). VLDB Endowment.

Du, R., Gkatzikis, L., Fischione, C., & Xiao, M. (2018). On maximizing sensor network lifetime by energy balancing. IEEE Transactions on Control of Network Systems, 5(3), 1206-1218.

Denantes, P., Bénézit, F., Thiran, P., & Vetterli, M. (2008). Which distributed averaging algorithm should i choose for my sensor network?. In IEEE INFOCOM 2008-The 27th Conference on Computer Communications (pp. 986-994). IEEE.

Dias, G. M., Bellalta, B., & Oechsner, S. (2016). A survey about prediction-based data reduction in wireless sensor networks. ACM Computing Surveys (CSUR), 49(3), 58.

Donoho, D. L. (2006). Compressed sensing. IEEE Transactions on information theory, 52(4), 1289-1306.

Duarte, M. F., Sarvotham, S., Wakin, M. B., Baron, D., &Baraniuk, R. G. (2005). Joint sparsity models for distributed compressed sensing. In Proceedings of the Workshop on Signal Processing with Adaptive Sparse Structured Representations. IEEE.

Di Lorenzo, P., & Sayed, A. H. (2012). Sparse distributed learning based on diffusion adaptation. IEEE Transactions on signal processing, 61(6), 1419-1433.

Duarte, M. F., Cevher, V., & Baraniuk, R. G. (2009). Model-based compressive sensing for signal ensembles. In 2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton) (pp. 244-250). IEEE.

Di Francesco, M., Das, S. K., & Anastasi, G. (2011). Data collection in wireless sensor networks with mobile elements: A survey. ACM Transactions on Sensor Networks (TOSN), 8(1), 7.

Donoho, D., & Kutyniok, G. (2013). Microlocal analysis of the geometric separation problem. Communications on Pure and Applied Mathematics, 66(1), 1-47.

Do, T. T., Gan, L., Nguyen, N. H., & Tran, T. D. (2012). Fast and efficient compressive sensing using structurally random matrices. IEEE Transactions on Signal Processing, 60(1), 139-154.

Dhimal, S., & Sharma, K. (2015). Energy conservation in wireless sensor networks by exploiting inter-node data similarity metrics. International Journal of Energy, Information and Communications, 6(2), 23-32.

Duarte, M. F., Shen, G., Ortega, A., & Baraniuk, R. G. (2012). Signal compression in wireless sensor networks. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 370(1958), 118-135.

Dang, T., Bulusu, N., & Feng, W. C. (2007). RIDA: A robust information-driven data compression architecture for irregular wireless sensor networks. In European Conference on Wireless Sensor Networks (pp. 133-149). Springer, Berlin, Heidelberg.

Ding, X., Chen, W., & Wassell, I. J. (2017). Joint sensing matrix and sparsifying dictionary optimization for tensor compressive sensing. IEEE Transactions on Signal Processing, 65(14), 3632-3646.

Dolas, P., & Ghosh, D. (2018). Distributed Compressive Data Gathering Framework for Correlated Data in Wireless Sensor Networks. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(1-6), 153-158.

Estrin, D., Girod, L., Pottie, G., & Srivastava, M. (2001). Instrumenting the world with wireless sensor networks. In icassp (Vol. 1, pp. 2033-2036).

Ee, C. T., & Bajcsy, R. (2004). Congestion control and fairness for many-to-one routing in sensor networks. In Proceedings of the 2nd international conference on Embedded networked sensor systems (pp. 148-161). ACM.

Ebrahimi, D., & Assi, C. (2015). On the interaction between scheduling and compressive data gathering in wireless sensor networks. IEEE Transactions on Wireless Communications, 15(4), 2845-2858.

Fan, Y., Chen, Q., & Yu, J. (2009). Topology control algorithm based on bottleneck node for large-scale WSNs. In 2009 International Conference on Computational Intelligence and Security (Vol. 1, p. 592-597). IEEE.

Fasolo, E., Rossi, M., Widmer, J., & Zorzi, M. (2007). In-network aggregation techniques for wireless sensor networks: a survey. IEEE Wireless Communications, 14(2), 70-87.

Felemban, E., Lee, C. G., & amp; Ekici, E. (2006). MMSPEED: multipath Multi-SPEED protocol for QoS guarantee of reliability and. Timeliness in wireless sensor networks. IEEE transactions on mobile computing, 5(6), 738-754.

Fazel, F., Fazel, M., & Stojanovic, M. (2013). Random access compressed sensing over fading and noisy communication channels. IEEE Transactions on Wireless Communications, 12(5), 2114-2125.

Felipeda Rocha Henriques, F., Lovisolo, L., & Barros da Silva, E. A. (2019). Rate-Distortion Performance and Incremental Transmission Scheme of Compressive Sensed Measurements in Wireless Sensor Networks. Sensors, 19(2), 266.

Fu, X., Yao, H., & Yang, Y. (2019). Modeling and analyzing cascading dynamics of the clustered wireless sensor network. Reliability Engineering & System Safety, 186, 1-10.

Firooz, M. H., & Roy, S. (2014). Link delay estimation via expander graphs. IEEE Transactions on Communications, 62(1), 170-181.

Gehrke, J., & Madden, S. (2004). Query processing in sensor networks. IEEE Pervasive computing, 3(1), 46-55.

Guestrin, C., Bodik, P., Thibaux, R., Paskin, M., & Madden, S. (2004). Distributed regression: an efficient framework for modeling sensor network data. In Proceedings of the 3rd international symposium on Information processing in sensor networks (pp. 1-10). ACM.

Galzarano, S., Fortino, G., & Liotta, A. (2014). A learning-based mac for energy efficient wireless sensor networks. In International Conference on Internet and Distributed Computing Systems (pp. 396-406). Springer, Cham.

Gupta, H., Navda, V., Das, S., & Chowdhary, V. (2008). Efficient gathering of correlated data in sensor networks. ACM Transactions on Sensor Networks (TOSN), 4(1), 4.

Gnawali, O., Jang, K. Y., Paek, J., Vieira, M., Govindan, R., Greenstein, B., & Kohler, E. (2006, October). The tenet architecture for tiered sensor networks. In Proceedings of the 4th international conference on Embedded networked sensor systems (pp. 153-166).

Ghaderi, M. R., Vakili, V. T., & Sheikhan, M. (2019). FGAF-CDG: fuzzy geographic routing protocol based on compressive data gathering in wireless sensor networks. Journal of Ambient Intelligence and Humanized Computing, 1-23.

Gan, H., Xiao, S., & Zhao, Y. (2018). A large class of chaotic sensing matrices for compressed sensing. Signal Processing, 149, 193-203.

Goela, N., &Gastpar, M. (2012). Reduced-dimension linear transform coding of correlated signals in networks. IEEE Transactions on Signal Processing, 60(6), 3174-3187.

Guo, D., Qu, X., Xiao, M., & Yao, Y. (2009). Comparative analysis on transform and reconstruction of compressed sensing in sensor networks. In 2009 WRI International Conference on Communications and Mobile Computing (Vol. 1, pp. 441-445). IEEE.

Heinzelman, W. R., Chandrakasan, A., &Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. In Proceedings of the 33rd annual Hawaii international conference on system sciences (pp. 10-pp). IEEE.

He, S., Chen, J., Li, X., Shen, X., & Sun, Y. (2011). Leveraging prediction to improve the coverage of wireless sensor networks. IEEE Transactions on Parallel and Distributed Systems, 23(4), 701-712.

Huang, J., & Soong, B. H. (2017). Hybrid compressive sensing for delay-efficient sustainable data gathering. In 2017 IEEE 85th Vehicular Technology Conference (VTC Spring) (pp. 1-5). IEEE.

Haupt, J., Castro, R., Nowak, R., Fudge, G., & Yeh, A. (2006). Compressive sampling for signal classification. In 2006 Fortieth Asilomar Conference on Signals, Systems and Computers (pp. 1430-1434). IEEE.

Haupt, J., Bajwa, W. U., Rabbat, M., & Nowak, R. (2008). Compressed sensing for networked data. IEEE Signal Processing Magazine, 25(2), 92-101.

Hempstead, M., Lyons, M. J., Brooks, D., & Wei, G. Y. (2008). Survey of hardware systems for wireless sensor networks. Journal of Low Power Electronics, 4(1), 11-20.

Haenggi, M. (2003). Energy-balancing strategies for wireless sensor networks. In IEEE International Symposium on Circuits and Systems (No. 4, pp. IV-828). IEEE; 1999.

Haenggi, M. (2004). Twelve reasons not to route over many short hops. In IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 (Vol. 5, pp. 3130-3134). IEEE.

Hoang, D. C., Yadav, P., Kumar, R., & Panda, S. K. (2013). Real-time implementation of a harmony search algorithm-based clustering protocol for energy-efficient wireless sensor networks. IEEE transactions on industrial informatics, 10(1), 774-783.

Huang, J., & Soong, B. H. (2019). Cost-Aware Stochastic Compressive Data Gathering for Wireless Sensor Networks. IEEE Transactions on Vehicular Technology, 68(2), 1525-1533.

Hooshmand, M., Rossi, M., Zordan, D., &Zorzi, M. (2015). Covariogram-based compressive sensing for environmental wireless sensor networks. IEEE Sensors Journal, 16(6), 1716-1729.

Intanagonwiwat, C., Govindan, R., Estrin, D., Heidemann, J., & Silva, F. (2003). Directed diffusion for wireless sensor networking. IEEE/ACM Transactions on Networking (ToN), 11(1), 2-16.

Intanagonwiwat, C., Estrin, D., Govindan, R., & Heidemann, J. (2002). Impact of network density on data aggregate on in wireless sensor networks. In ICDCS (Vol. 2, p. 457).

Idrees, A. K., Deschinkel, K., Salomon, M., & Couturier, R. (2015). Distributed lifetime coverage optimization protocol in wireless sensor networks. The journal of supercomputing, 71(12), 4578-4593.

Jiang, R., & Chen, B. (2005). Fusion of censored decisions in wireless sensor networks. IEEE Transactions on Wireless Communications, 4(6), 2668-2673.

Jones, C. E., Sivalingam, K. M., Agrawal, P., & Chen, J. C. (2001). A survey of energy efficient network protocols for wireless networks. Wireless networks, 7(4), 343-358.

Ju, X., Liu, W., Zhang, C., Liu, A., Wang, T., Xiong, N. N., & Cai, Z. (2018). An energy conserving and transmission radius adaptive scheme to optimize performance of energy harvesting sensor networks. Sensors, 18(9), 2885.

Jarry, A., Leone, P., Powell, O., & Rolim, J. (2006). An optimal data propagation algorithm for maximizing the lifespan of sensor networks. In International Conference on Distributed Computing in Sensor Systems (pp. 405-421). Springer, Berlin, Heidelberg.

Jain, N., Bohara, V. A., & Gupta, A. (2019). iDEG: Integrated Data and Energy Gathering Framework for Practical Wireless Sensor Networks Using Compressive Sensing. IEEE Sensors Journal, 19(3), 1040-1051.

Krishnamachari, B., Estrin, D., & Wicker, S. (2002). Modelling data- centric routing in wireless sensor networks. In IEEE infocom (Vol. 2, pp. 39-44).

Khojastepour, M. A., &Aazhang, B. (2004). The capacity of average and peak power constrained fading channels with channel side information. In 2004 IEEE Wireless Communications and Networking Conference (IEEE Cat. No. 04TH8733) (Vol. 1, pp. 77-82). IEEE.

Kumar, K., & Lu, Y. H. (2010). Cloud computing for mobile users: Can off-loading computation save energy? Computer, (4), 51-56.

Karakus, C., Gurbuz, A. C., & Tavli, B. (2013). Analysis of energy efficiency of compressive sensing in wireless sensor networks. IEEE Sensors Journal, 13(5),1999-2008.

Kuila, P., & Jana, P. K. (2014). Approximation schemes for load balanced clustering in wireless sensor networks. The Journal of Supercomputing, 68(1), 87-105.

Kong, L., Xia, M., Liu, X. Y., Chen, G., Gu, Y., Wu, M. Y., & Liu, X. (2014). Data loss and reconstruction in wireless sensor networks. IEEE Transactions on Parallel and Distributed Systems, 25(11), 2818-2828.

Kulkarni, R. V., Forster, A., & Venayagamoorthy, G. K. (2010). Computational intelligence in wireless sensor networks: A survey. IEEE communications surveys & tutorials, 13(1), 68-96.

Kumar, D. P., Amgoth, T., & Annavarapu, C. S. R. (2019). Machine learning algorithms for wireless sensor networks: A survey. Information Fusion, 49, 1-25.

Kang, L., Chen, R. S., Chen, Y. C., Wang, C. C., Li, X., & Wu, T. Y. (2019). Using Cache Optimization Method to Reduce Network Traffic in Communication Systems Based on Cloud Computing. IEEE Access. Pg 124397 – 124409.

Luo, H., Liu, Y., & Das, S. K. (2007). Routing correlated data in wireless sensor networks: A survey. IEEE Network, 21(6).

Lin, Y., Chen, B., & Varshney, P. K. (2005). Decision fusion rules in multi-hop wireless sensor networks. IEEE Transactions on Aerospace and Electronic Systems, 41(2), 475-488.

Lu, M., Wu, J., Cardei, M., & Li, M. (2009). Energy-efficient connected coverage of discrete targets in wireless sensor networks. International Journal of Ad Hoc and Ubiquitous Computing, 4(3-4), 137-147.

Lee, J. J., Krishnamachari, B., & Kuo, C. C. J. (2008). Aging analysis in large-scale wireless sensor networks. Ad Hoc Networks, 6(7), 1117-1133.

Laska, J. N., Boufounos, P. T., Davenport, M. A., & Baraniuk, R. G. (2011). Democracy in action: Quantization, saturation, and compressive sensing. Applied and Computational Harmonic Analysis, 31(3), 429-443.

Liu, Z., Li, Z., Li, M., Xing, W., Lu, D., Liu, Z., & Lu, D. (2016). Path reconstruction in dynamic wireless sensor networks using compressive sensing. IEEE/ACM Transactions on Networking (TON), 24(4), 1948-1960.

Luo, J., & Hubaux, J. P. (2005). Joint mobility and routing for lifetime elongation in wireless sensor networks. In Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies. (Vol. 3, pp. 1735-1746). IEEE.

Liu, A. F., Wu, X. Y., Chen, Z. G., & Gui, W. H. (2010). Research on the energy hole problem based on unequal cluster-radius for wireless sensor networks. Computer communications, 33(3), 302-321.

Lv, C., Wang, Q., Yan, W., & Li, J. (2019). Compressive Sensing-based sequential data gathering in WSNs. Computer Networks. Vol 154 pp47-59.

Li, C., Ye, M., Chen, G., & Wu, J. (2005). An energy-efficient unequal clustering mechanism for wireless sensor networks. In IEEE International Conference on Mobile Adhoc and Sensor Systems Conference, 2005. (pp. 8-pp). IEEE.

Ling, Q., & Tian, Z. (2010). Decentralized sparse signal recovery for compressive sleeping wireless sensor networks. IEEE Transactions on Signal Processing, 58(7), 3816-3827.

Luo, C., Wu, F., Sun, J., & Chen, C. W. (2009). Compressive data gathering for large-scale wireless sensor networks. In Proceedings of the 15th annual international conference on Mobile computing and networking (pp. 145-156). ACM.

Li, Y., & Parker, L. E. (2014). Nearest neighbor imputation using spatial–temporal correlations in wireless sensor networks. Information Fusion, 15, 64-79.

Lu, H., & Bo, L. (2019). WDL Recon Net: Compressive Sensing Reconstruction with Deep Learning over Wireless Fading Channels. IEEE Access. Volume 7 Pg 24440-24451.DOI 10.1109/ACCESS.2019.2900715.

Li, Z., Liu, Y., Ma, M., Liu, A., Zhang, X., & Luo, G. (2018). MSDG: A novel green data gathering scheme for wireless sensor networks. Computer Networks, 142, 223-239.

Li, Y., & Liang, Y. (2018). Compressed Sensing in Multi-Hop Large-Scale Wireless Sensor Networks Based on Routing Topology Tomography. IEEE Access, 6, 27637-27650.

Liu, J., Huang, K., & Yao, X. (2018). Common-Innovation Subspace Pursuit for Distributed Compressed Sensing in Wireless Sensor Networks. IEEE Sensors Journal, 19(3), 1091-1103.

Liu, C., Wu, K., & Tsao, M. (2005). Energy efficient information collection with the ARIMA model in wireless sensor networks. In GLOBECOM'05. IEEE Global Telecommunications Conference, 2005. (Vol. 5, pp. 5-pp).

Mahfoudh, S., & Minet, P. (2008). Survey of energy efficient strategies in wireless ad hoc and sensor networks. In Seventh International Conference on Networking (icn 2008) (pp. 1-7). IEEE.

Mahyar, H., Hasheminezhad, R., Ghalebi, E., Nazemian, A., Grosu, R., Movaghar, A., & Rabiee, H. R. (2018). Compressive sensing of high betweenness centrality nodes in networks. Physica A: Statistical Mechanics and its Applications, 497, 166-184.

Madden, S., Szewczyk, R., Franklin, M. J., & Culler, D. (2002). Supporting aggregate queries over ad-hoc wireless sensor networks. In Proceedings Fourth IEEE Workshop on Mobile Computing Systems and Applications (pp. 49-58). IEEE.

Msechu, E. J., & Giannakis, G. B. (2011). Sensor-centric data reduction for estimation with WSNs via censoring and quantization. IEEE Transactions on Signal Processing, 60(1), 400-414.

Mahmood, A., Shi, K., Khatoon, S., & Xiao, M. (2013). Data mining techniques for wireless sensor networks: A survey. International Journal of Distributed Sensor Networks, 9(7), 406316.

Marvasti, F., Amini, A., Haddadi, F., Soltanolkotabi, M., Khalaj, B. H., Aldroubi, A., & Chambers, J. (2012). A unified approach to sparse signal processing. EURASIP journal on advances in signal processing, 2012(1), 44.

Mishali, M., & Eldar, Y. C. (2012). Xampling: Compressed sensing for analog signals. In Compressed Sensing: Theory and Applications (No. 3). Cambridge University Press. Chapter 3 Pg. 88-147.

Mascolo, C., Hailes, S., Lymberopoulos, L., Picco, G. P., Costa, P., Blair, G., ... & Rónai, M. A. (2005). Survey of middleware for networked embedded systems. Project Report: http://www. ist-runes. org/docs/deliverables/D5_01. pdf.

Marco, D., Duarte-Melo, E. J., Liu, M., & Neuhoff, D. L. (2003). On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data. In Information Processing in Sensor Networks (pp. 1-16). Springer, Berlin, Heidelberg.

Meng, J., Li, H., & Han, Z. (2009). Sparse event detection in wireless sensor networks using compressive sensing. In 2009 43rd Annual Conference on Information Sciences and Systems (pp. 181-185). IEEE.

Masiero, R., Quer, G., Munaretto, D., Rossi, M., Widmer, J., & Zorzi, M. (2009). Data acquisition through joint compressive sensing and principal component analysis. In GLOBECOM 2009-2009 IEEE Global Telecommunications Conference (pp. 1-6). IEEE.

Mehrjoo, S., & Khunjush, F. (2018). Accurate compressive data gathering in wireless sensor networks using weighted spatio-temporal compressive sensing. Telecommunication Systems, 68(1), 79-88.

Masoum, A., Meratnia, N., & Havinga, P. (2018). Coalition Formation Based Compressive Sensing in Wireless Sensor Networks. Sensors, 18(7), 2331.

Ma, X., Liang, J., Liu, R., Ni, W., Li, Y., Li, R., & Qi, C. (2018). A survey on data storage and information discovery in the WSANs-based edge computing systems. Sensors, 18(2), 546.

Nguyen, H. X., & Thiran, P. (2006). Using end-to-end data to infer lossy links in sensor networks. In IEEE Infocom 2006 (No. CONF).

Nguyen, T. L., & Shin, Y. (2013). Deterministic sensing matrices in compressive sensing: a survey. The Scientific World Journal, 2013, 1-6.http://dx.doi.org/10.1155/2013/192795.

Nguyen, M. T., Teague, K. A., & Rahnavard, N. (2016). CCS: Energy-efficient data collection in clustered wireless sensor networks utilizing block-wise compressive sensing. Computer Networks, 106, 171-185.

Nayak, A., & Stojmenovic, I. (2010). Wireless sensor and actuator networks: algorithms and protocols for scalable coordination and data communication. John Wiley & Sons. Chapter 6.

Nie, L., Wang, X., Wan, L., Yu, S., Song, H., & Jiang, D. (2018). Network traffic prediction based on deep belief network and spatiotemporal compressive sensing in wireless mesh backbone networks. Wireless Communications and Mobile Computing, 2018. https://doi.org/10.1155/2018/1260860.

Pattem, S., Krishnamachari, B., & Govindan, R. (2008). The impact of spatial correlation on routing with compression in wireless sensor networks. ACM Transactions on Sensor Networks (TOSN), 4(4), 24.

Pradhan, S. S., Kusuma, J., & Ramchandran, K. (2002). Distributed compression in a dense microsensor network. IEEE Signal Processing Magazine, 19(2), 51-60.

Polastre, J., Hill, J., & Culler, D. (2004). Versatile low power media access for wireless sensor networks. In Proceedings of the 2nd international conference on Embedded networked sensor systems (pp. 95-107). ACM.

Perillo, M., Cheng, Z., & Heinzelman, W. (2004). On the problem of unbalanced load distribution in wireless sensor networks. In IEEE Global Telecommunications Conference Workshops, 2004. GlobeCom Workshops 2004. (pp. 74-79). IEEE.

Puccinelli, D., & Haenggi, M. (2008). Arbutus: Network-layer load balancing for wireless sensor networks. In 2008 IEEE Wireless Communications and Networking Conference (pp. 2063-2068). IEEE.

Patterson, S., Eldar, Y. C., & Keidar, I. (2014). Distributed compressed sensing for static and time-varying networks. IEEE Transactions on Signal Processing, 62(19), 4931-4946.

Quan, L., Xiao, S., Xue, X., & Lu, C. (2016). Neighbor-aided spatial-temporal compressive data gathering in wireless sensor networks. IEEE Communications Letters, 20(3), 578-581.

Quer, G., Masiero, R., Pillonetto, G., Rossi, M., & Zorzi, M. (2012). Sensing, compression, and recovery for WSNs: Sparse signal modeling and monitoring framework. IEEE Transactions on Wireless Communications, 11(10), 3447-3461.

Qin, Z., Fan, J., Liu, Y., Gao, Y., & Li, G. Y. (2018). Sparse representation for wireless communications: A compressive sensing approach. IEEE Signal Processing Magazine, 35(3), 40-58.

Qiao, J., & Zhang, X. (2018). Compressive Data Gathering Based on Even Clustering for Wireless Sensor Networks. IEEE ACCESS, 6, 24391-24410.

Romer, K., & Renner, B. C. (2008). Aggregating sensor data from overlapping multi-hop network neighborhoods: Push or pull?. In 2008 5th International Conference on Networked Sensing Systems (pp. 107-110). IEEE.

Raghunathan, V., Schurgers, C., Park, S., & Srivastava, M. B. (2002). Energy-aware wireless microsensor networks. IEEE Signal processing magazine, 19(2), 40-50.

Raza, U., Camerra, A., Murphy, A. L., Palpanas, T., & Picco, G. P. (2012). What does model-driven data acquisition really achieve in wireless sensor networks?. In 2012 IEEE International Conference on Pervasive Computing and Communications (pp. 85-94). IEEE.

Raza, U., Camerra, A., Murphy, A. L., Palpanas, T., & Picco, G. P. (2015). Practical data prediction for real-world wireless sensor networks. IEEE Transactions on Knowledge and Data Engineering, 27(8), 2231-2244.

Rago, C., Willett, P., & Bar-Shalom, Y. (1996). Censoring sensors: A low-communication-rate scheme for distributed detection. IEEE Transactions on Aerospace and Electronic Systems, 32(2), 554-568.

Raghunathan, V., Schurgers, C., Park, S., & Srivastava, M. B. (2002). Energy-aware wireless microsensor networks. IEEE Signal processing magazine, 19(2), 40-50.

Rault, T., Bouabdallah, A., & Challal, Y. (2014). Energy efficiency in wireless sensor networks: A top-down survey. Computer Networks, 67, 104-122.

Rao, Y., Zhao, G., Wang, W., Zhang, J., Jiang, Z., & Wang, R. (2019). Adaptive Data Acquisition with Energy Efficiency and Critical-Sensing Guarantee for Wireless Sensor Networks. Sensors, 19(12), 2654.

Rajaram, S., Karuppiah, A. B., & Kumar, K. V. (2014). Secure routing path using trust values for wireless sensor networks. International Journal on Cryptography and Information Security (IJCIS), Vol. 4, No. 2, 27-35.

Razzaque, M. A., Bleakley, C., & Dobson, S. (2013). Compression in wireless sensor networks: A survey and comparative evaluation. ACM Transactions on Sensor Networks (TOSN), 10(1), 5.

Singh, V. K., Kumar, M., & Verma, S. (2018). Node scheduling and compressed sampling for event reporting in WSNs. IEEE Transactions on Network Science and Engineering. pp. 1–1, 2018.

Shih, E., Calhoun, B. H., Cho, S. H., & Chandrakasan, A. P. (2001). Energy-efficient link layer for wireless microsensor networks. In Proceedings IEEE Computer Society Workshop on VLSI 2001. Emerging Technologies for VLSI Systems (pp. 16-21). IEEE.

Sherazi, H. H. R., Grieco, L. A., & Boggia, G. (2018). A comprehensive review on energy harvesting MAC protocols in WSNs: Challenges and tradeoffs. Ad Hoc Networks, 71, 117-134.

Shah, R. C., & Rabaey, J. M. (2002). Energy aware routing for low energy ad hoc sensor networks. In 2002 IEEE Wireless Communications and Networking Conference Record. WCNC 2002 (Cat. No. 02TH8609) (Vol. 1, pp. 350-355). IEEE.

Sarvotham, S., Baron, D., Wakin, M., Duarte, M. F., & Baraniuk, R. G. (2005). Distributed compressed sensing of jointly sparse signals. In Asilomar conference on signals, systems, and computers (pp. 1537-1541).

Singh, V. K., & Kumar, M. (2018). A compressed sensing approach to resolve the energy hole problem in large scale WSNs. Wireless Personal Communications, 99(1), 185-201.

Singh, V. K., Verma, S., & Kumar, M. (2019). ODECS: An On-Demand Explosion-Based Compressed Sensing Using Random Walks in Wireless Sensor Networks. IEEE Systems Journal. 1 – 10.

Sun, Z., Tao, R., Xiong, N., & Pan, X. (2018). CS-PLM: Compressive sensing data gathering algorithm based on packet loss matching in sensor networks. Wireless Communications and Mobile Computing, 2018.

Shokouhifar, M., &Jalali, A. (2015). A new evolutionary based application specific routing protocol for clustered wireless sensor networks. AEU-International Journal of Electronics and Communications, 69(1), 432-441.

Shen, Y., Hu, W., Rana, R., & Chou, C. T. (2013). Nonuniform compressive sensing for heterogeneous wireless sensor networks. IEEE Sensors journal, 13(6), 2120-2128.

Shen, G., Narang, S. K., & Ortega, A. (2009). Adaptive distributed transforms for irregularly sampled wireless sensor networks. In 2009 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 2225-2228). IEEE.

Shen, G., & Ortega, A. (2010). Transform-based distributed data gathering. IEEE Transactions on Signal Processing, 58(7), 3802-3815.

Shekaramiz, M., Moon, T. K., & Gunther, J. H. (2019). Bayesian Compressive Sensing of Sparse Signals with Unknown Clustering Patterns. Entropy, 21(3), 247.

Tilak, S., Abu-Ghazaleh, N. B., & Heinzelman, W. (2002). A taxonomy of wireless micro-sensor network models. ACM SIGMOBILE Mobile Computing and Communications Review, 6(2), 28-36.

Tropp, J. A., Gilbert, A. C., & Strauss, M. J. (2006). Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit. Signal processing, 86(3), 572-588.

Talari, A., & Rahnavard, N. (2016). CStorage: Decentralized compressive data storage in wireless sensor networks. Ad Hoc Networks, 37, 475-485.

Thakkar, A., & Kotecha, K. (2014). Cluster head election for energy and delay constraint applications of wireless sensor network. IEEE sensors Journal, 14(8), 2658-2664.

Tayeh, G. B., Makhoul, A., Perera, C., & Demerjian, J. (2019). A Spatial-Temporal Correlation Approach for Data Reduction in Cluster-Based Sensor Networks. IEEE Access, 7, 50669-50680.

Tulone, D., & Madden, S. (2006). PAQ: Time series forecasting for approximate query answering in sensor networks. In European Workshop on Wireless Sensor Networks (pp. 21-37). Springer, Berlin, Heidelberg.

Vuran, M. C., Akan, Ö. B., &Akyildiz, I. F. (2004). Spatio-temporal correlation: theory and applications for wireless sensor networks. Computer Networks, 45(3), 245-259.

Vlajic, N., & Xia, D. (2006). Wireless sensor networks: to cluster or not to cluster?. In Proceedings of the 2006 International Symposium on on World of Wireless, Mobile and Multimedia Networks (pp. 258-268). IEEE Computer Society.

Vuran, M. C., & Akyildiz, I. F. (2006). Spatial correlation-based collaborative medium access control in wireless sensor networks. IEEE/ACM Transactions on Networking, 14(2), 316-329.

Vardi, Y. (1996). Network tomography: Estimating source-destination traffic intensities from link data. Journal of the American statistical association, 91(433), 365-377.

Van Nguyen, T., Quek, T. Q., & Shin, H. (2018). Joint channel identification and estimation in wireless network: Sparsity and optimization. IEEE Transactions on Wireless Communications, 17(5), 3141-3153.

Wang, F., & Liu, J. (2010). Networked wireless sensor data collection: issues, challenges, and approaches. IEEE Communications Surveys & Tutorials, 13(4), 673-687.

Wagner, R. S., Baraniuk, R. G., Du, S., Johnson, D. B., & Cohen, A. (2006). An architecture for distributed wavelet analysis and processing in sensor networks. In Proceedings of the 5th international conference on Information processing in sensor networks (pp. 243-250). ACM.

Wang, Y. C., & Wei, C. T. (2016). Lightweight, latency aware routing for data compression in wireless sensor networks with heterogeneous traffics. Wireless Communications and Mobile Computing, 16(9), 1035-1049.

Wang, Q., Hempstead, M., & Yang, W. (2006). A realistic power consumption model for wireless sensor network devices. In 2006 3rd annual IEEE communications society on sensor and ad hoc communications and networks (Vol. 1, pp. 286-295). IEEE.

Wang, W., Wainwright, M. J., & Ramchandran, K. (2010). Information-theoretic limits on sparse signal recovery: Dense versus sparse measurement matrices. IEEE Transactions on Information Theory, 56(6), 2967-2979.

Wang, W., Garofalakis, M., & amp; Ramchandran, K. (2007). Distributed sparse random projections for refinable approximation. In Proceedings of the 6th international conference on Information processing in sensor networks (pp. 331-339). ACM.

Wang, J., Tang, S., Yin, B., & Li, X. Y. (2012). Data gathering in wireless sensor networks through intelligent compressive sensing. In 2012 Proceedings IEEE INFOCOM (pp. 603-611). IEEE.

Wu, J. Y., Yang, M. H., & Wang, T. Y. (2018). Energy-efficient sensor censoring for compressive distributed sparse signal recovery. IEEE Transactions on Communications, 66(5), 2137-2152.

Wang, X., Zhou, Q., & Cheng, C. T. (2019). A UAV-assisted topology-aware data aggregation protocol in WSN. Physical Communication, 34, 48-57.

Wang, X., Zhou, Q., & Tong, J. (2019) b. V-Matrix-Based Scalable Data Aggregation Scheme in WSN. IEEE Access, 7, 56081-56094.

Wang, M. M., Cao, J. N., Li, J., & Dasi, S. K. (2008). Middleware for wireless sensor networks: A survey. Journal of computer science and technology, 23(3), 305-326.

Wu, X., & Chen, G. (2007). Dual-sink: using mobile and static sinks for lifetime improvement in wireless sensor networks. In 2007 16th International Conference on Computer Communications and Networks (pp. 1297-1302). IEEE.

Willig, A., Matheus, K., & Wolisz, A. (2005). Wireless technology in industrial networks. Proceedings of the IEEE, 93(6), 1130-1151.

Wu, X., Xiong, Y., Yang, P., Wan, S., & Huang, W. (2014). Sparsest random scheduling for compressive data gathering in wireless sensor networks. IEEE Transactions on Wireless Communications, 13(10), 5867-5877.

Xu, X., Ansari, R., Khokhar, A., & Vasilakos, A. V. (2015). Hierarchical data aggregation using compressive sensing (HDACS) in WSNs. ACM Transactions on Sensor Networks (TOSN), 11(3), 45.

Xiang, L., Luo, J., & Vasilakos, A. V. (2011). Compressed data aggregation for energy efficient wireless sensor networks. In SECON (Vol. 2011, pp. 46-54).

Xie, R., &Jia, X. (2013). Transmission-efficient clustering method for wireless sensor networks using compressive sensing. IEEE transactions on parallel and distributed systems. 25(3) 806-815.

Xu, L., Hao, X., Lane, N. D., Liu, X., & Moscibroda, T. (2015). Cost-aware compressive sensing for networked sensing systems. In Proceedings of the 14th international conference on Information Processing in Sensor Networks (pp. 130-141). ACM.

Xu, W., Mallada, E., & Tang, A. (2011). Compressive sensing over graphs. In 2011 Proceedings IEEE INFOCOM (pp. 2087-2095). IEEE.

Xu, S., de Lamare, R. C., & Poor, H. V. (2015). Distributed compressed estimation based on compressive sensing. IEEE Signal Processing Letters, 22(9), 1311-1315.

Xia, J., Zhou, F., Lai, X., Zhang, H., Chen, H., Yang, Q., & Zhao, J. (2018). Cache aided decode-and-forward relaying networks: From the spatial view. Wireless Communications and Mobile Computing, 2018.

Xu, Y., Sun, G., Geng, T., & He, J. (2019). Low-Energy Data Collection in Wireless Sensor Networks Based on Matrix Completion. Sensors, 19(4), 945.

Xu, J., Liu, W., Lang, F., Zhang, Y., & Wang, C. (2010). Distance measurement model based on RSSI in WSN. Wireless Sensor Network, 2(08), 606.

Xiang, L., Luo, J., & Rosenberg, C. (2013). Compressed data aggregation: Energy-efficient and high-fidelity data collection. IEEE/ACM Transactions on Networking (TON), 21(6), 1722-1735.

Yang, Q., He, S., Li, J., Chen, J., & Sun, Y. (2014). Energy-efficient probabilistic area coverage in wireless sensor networks. IEEE Transactions on vehicular technology, 64(1), 367-377.

Yoon, S., & Shahabi, C. (2005). Exploiting spatial correlation towards an energy efficient clustered aggregation technique (cag)[wireless sensor network applications]. In IEEE International Conference on Communications, 2005. ICC 2005. 2005 (Vol. 5, pp. 3307-3313). IEEE.

Yoon, S., &Shahabi, C. (2007). The Clustered AGgregation (CAG) technique leveraging spatial and temporal correlations in wireless sensor networks. ACM Transactions on Sensor Networks (TOSN), 3(1), 3.

Yick, J., Mukherjee, B., & Ghosal, D. (2008). Wireless sensor network survey. Computer networks, 52(12), 2292-2330.

Yamasaki, K., & Ohtsuki, T. (2005). Design of energy-efficient wireless sensor networks with censoring, on-off, and censoring and on-off sensors based on mutual information. In 2005 IEEE 61st Vehicular Technology Conference (Vol. 2, pp. 1312-1316). IEEE.

Yang, A. Y., Gastpar, M., Bajcsy, R., & Sastry, S. S. (2010). Distributed sensor perception via sparse representation. Proceedings of the IEEE, 98(6), 1077-1088.

Yang, H., & Wang, X. (2018). ECOCS: Energy consumption optimized compressive sensing in group sensor networks. Computer Networks, 146, 159-166.

Yuan, F., Zhan, Y., & Wang, Y. (2013). Data density correlation degree clustering method for data aggregation in WSN. IEEE Sensors Journal, 14(4), 1089-1098.

Yang, X., Tao, X., Dutkiewicz, E., Huang, X., Guo, Y. J., & Cui, Q. (2013). Energy-efficient distributed data storage for wireless sensor networks based on compressed sensing and network coding. IEEE Transactions on Wireless Communications, 12(10), 5087-5099.

Zhang, H., & Hou, J. C. (2005). Maintaining sensing coverage and connectivity in large sensor networks. Ad Hoc & Sensor Wireless Networks, 1(1-2), 89-124.

Zhang, H., & Shen, H. (2008). Balancing energy consumption to maximize network lifetime in data-gathering sensor networks. IEEE Transactions on parallel and distributed systems, 20(10), 1526-1539.

Zhang, P., Chen, C., & Liu, M. (2009). The application of compressed sensing in wireless sensor network. In 2009 International Conference on Wireless Communications & Signal Processing (pp. 1-5). IEEE.

Zhang, B., Cheng, X., Zhang, N., Cui, Y., Li, Y., & Liang, Q. (2011). Sparse target counting and localization in sensor networks based on compressive sensing. In 2011 Proceedings IEEE INFOCOM (pp. 2255-2263). IEEE.

Zaeemzadeh, A., Joneidi, M., &Rahnavard, N. (2017). Adaptive non-uniform compressive sampling for time-varying signals. In 2017 51st Annual Conference on Information Sciences and Systems (CISS) (pp. 1-6). IEEE.

Zheng, H., Xiao, S., Wang, X., Tian, X., & Guizani, M. (2013). Capacity and delay analysis for data gathering with compressive sensing in wireless sensor networks. IEEE Transactions on Wireless Communications, 12(2), 917-927.

Zhang, D. G., Zhang, T., Zhang, J., Dong, Y., & Zhang, X. D. (2018). A kind of effective data aggregating method based on compressive sensing for wireless sensor network. EURASIP Journal on Wireless Communications and Networking, 2018(1), 159.

Zoubir, A. M., & Iskandler, D. R. (2007). Bootstrap methods and applications. IEEE Signal Processing Magazine, 24(4), 10-19.

Zhang, P., Wang, J., &Guo, K. (2018). Compressive sensing and random walk-based data collection in wireless sensor networks. Computer Communications, 129, 43-53.

Zhou, X., Ji, X., Chen, Y. C., Li, X., & Xu, W. (2018). LESS: Link Estimation with Sparse Sampling in Intertidal WSNs. Sensors, 18(3), 747.

Zhang, L., Niu, D., Song, E., & Shi, Q. (2019). Closed-Form Solution for Optimal Compression Matrix Design in Distributed Estimation. IEEE Access, 7, 5045-5056.

Zhou, Z., Li, C., Yang, S., & Guang, X. (2019). A Practical Inner Codes for BATS Codes in Wireless Multi-hop Networks. IEEE Transactions on Vehicular Technology, 68(3), 2751-2762.

Zhou, Y., Yang, L., Yang, L., & Ni, M. (2019). Novel Energy-Efficient Data Gathering Scheme Exploiting Spatial-Temporal Correlation for Wireless Sensor Networks. Wireless Communications and Mobile Computing, 2019.

Zuhairy, R. M., & Al Zamil, M. G. (2018). Energy-efficient load balancing in wireless sensor network: An application of multinomial regression analysis. International Journal of Distributed Sensor Networks, 14(3), 1550147718764641.

SUPPLEMENTARY

LIST OF PUBLICATIONS AND PAPERS PRESENTED

[1] Alagirisamy, M., & Chow, C. O. (2018). An energy-based cluster head selection unequal clustering algorithm with dual sink (ECH-DUAL) for continuous monitoring applications in wireless sensor networks. Cluster Computing, 21(1), 91-103.

[2] Alagirisamy, M., Chow, C. O., & Noordin, K. A. B. (2018). Dual Sink Based-Energy Aware Unequal Clustering Routing Algorithm for Continuous Monitoring Applications in Wireless Sensor Network. In 2018 IEEE Conference on Open Systems (ICOS) (pp. 47-53). IEEE.

[3] Mukil Alagirisamy, Chee-Onn Chow, Kamarul Ariffin Bin Noordin, (2019). Compressive sensing with perceptron-based routing for varying traffic intensity based on capsule networks, Computers & Electrical Engineering, 79,106446, ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2019.08.008.

[4] Alagirisamy, M., Chow, C. & Noordin, K.A.B. Intelligence Framework Based Analysis of Spatial–Temporal Data with Compressive Sensing Using Wireless Sensor Networks, Wireless Personal Communication 112, 91–103 (2020). https://doi.org/10.1007/s11277-019-07017-2.