VISVESVARAYA TECHNOLOGICAL UNIVERSITY

Jnana Sangama, Belgaum-590018

A PROJECT REPORT **(15CSP85)** ON

"A NEW ENERGY EFFICIENT TRANSMISSION SCHEME FOR WIRELESS SENSOR NETOWRKS"

Submitted in Partial fulfillment of the Requirements for the Degree of

Bachelor of Engineering in Computer Science & Engineering By

ADITYA SATHISH (1CR16CS010)

ACHUTH S (1CR16CS007)

Under the Guidance of,

MR. PRATHAM MAJUMDER

Assistant Professor, Dept. of CSE

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CMR INSTITUTE OF TECHNOLOGY

#132, AECS LAYOUT, IT PARK ROAD, KUNDALAHALLI, BANGALORE-560037

CMR INSTITUTE OF TECHNOLOGY

#132, AECS LAYOUT, IT PARK ROAD, KUNDALAHALLI, BANGALORE-560037

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

Certified that the project work entitled **"A NEW ENERGY EFFICIENT TRANSMISSION SCHEME FOR WIRELESS SENSOR NETWORKS"** carried out by **Mr. ADITYA SATHISH** USN **1CR16CS010, Mr. ACHUTH S USN ICR16CS007,** bonafide students of CMR Institute of Technology, in partial fulfillment for the award of **Bachelor of Engineering** in Computer Science and Engineering of the Visveswaraiah Technological University, Belgaum during the year 2019-2020. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the Report deposited in the departmental library.

The project report has been approved as it satisfies the academic requirements in respect of Project work prescribed for the said Degree.

__________________	__________________	__________________
Mr. Pratham M	**Dr. Prem Kumar Ramesh**	**Dr. Sanjay Jain**
Assistant Professor	**Professor & Head**	**Principal**
Dept. of CSE, CMRIT	**Dept. of CSE, CMRIT**	**CMRIT**

External Viva

Name of the examiners	Signature with date
1.	__________________
2.	__________________

DECLARATION

We, the students of Computer Science and Engineering, CMR Institute of Technology, Bangalore declare that the work entitled " **A NEW ENERGY EFFICIENT TRANSMISSION SCHEME FOR WIRELESS SENSOR NETWORKS** " has been successfully completed under the guidance of Prof. Pratham M, Computer Science and Engineering Department, CMR Institute of technology, Bangalore. This dissertation work is submitted in partial fulfillment of the requirements for the award of Degree of Bachelor of Engineering in Computer Science and Engineering during the academic year 2019 - 2020. Further the matter embodied in the project report has not been submitted previously by anybody for the award of any degree or diploma to any university.

Place: Bangalore

Date: 22/06/2020

Team members:

ADITYA SATHISH (1CR16CS010) ___________________

ACHUTH S (1CR16CS007) ___________________

ABSTRACT

An energy-efficient communication scheme for low-power wireless networks is proposed which is based on encoding the messages by a novel source coding technique, called Dual-message Compression with Variable Null Symbol (DCV NS), followed by using the concept of Silent Communication. The source coding technique used in this scheme interleaves bits from two messages and encodes every two consecutive bits of this interleaved dual-message by one of four symbol values so that the overall message length is reduced by a factor of two. The transmitter is kept silent during the periods of the most dominant symbol which is dynamically changed. A hybrid modulation / demodulation technique using non-coherent F SK and ASK is used to have 43.46% energy savings on an average at the transmitter using a channel with Additive White Gaussian Noise (AW GN), and 50% energy savings at the receiver over the conventional Binary Frequency Shift Keying (BF SK), with equal likelihood of all possible strings of binary messages. Simulation results with real-life sensor data show that we can get still higher energy savings at the transmitter, ranging from about 54.1% to 58.9%. Also, the transmission time is reduced by a factor of two due to the reduced length of the encoded message. The scheme is suitable for multihop communication in low cost wireless sensor network applications.

ACKNOWLEDGEMENT

We take this opportunity to express our sincere gratitude and respect to **CMR Institute of Technology, Bengaluru** for providing us a platform to pursue our studies and carry out our final year project

We have a great pleasure in expressing our deep sense of gratitude to **Dr. Sanjay Jain,** Principal, CMRIT, Bangalore, for his constant encouragement.

We would like to thank **Dr. Prem Kumar Ramesh,** Professor and Head, Department of Computer Science and Engineering, CMRIT, Bangalore, who has been a constant support and encouragement throughout the course of this project.

We consider it a privilege and honor to express our sincere gratitude to our guide **Mr. Pratham Majumder, Assistant Professor ,** Department of Computer Science and Engineering, for the valuable guidance throughout the tenure of this review.

We also extend our thanks to all the faculty of Computer Science and Engineering who directly or indirectly encouraged us.

Finally, we would like to thank our parents and friends for all their moral support they have given us during the completion of this work.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In recent years, there has been a phenomenal growth in the use of wireless communication in almost every sphere of life. Many real-life applications are heavily dependent on various types of sensors forming wireless sensor networks (WSNs). It is expected that around 2020, about 30 billion devices will be interconnected through such WSNs forming the so-called Internet of Things (IoT). The devices so connected are often powered by battery of limited lifetime. Hence, energy-efficient communication is of paramount importance in the WSNs. Energy can be conserved in various layers of communication protocol which include reducing energy consumption through avoiding/reducing collisions and packet retransmissions in the MAC layer [1]–[3], intelligent routing techniques [4] in the network layer, and appropriate aggregation of data from different sensors in the application layer. In the physical layer as well, we can reduce energy consumption in the following way. Instead of spending energy on every transmitted bit, whether ’0’ or ’1’ as in the conventional Energy-based Transmission (EbT) schemes, we can selectively choose some bit periods during which the transmitter does not spend any energy. In fact, Zhu and Shivakumar first proposed [5] such a paradigm known as Communication through Silence (CtS) which involves the use of silent periods of transmitter when it does not transmit any bit at all. CtS, however, suffers from the disadvantage of requiring a communication time exponential in the number of symbols to be transmitted which can be reduced to some extent by an alternative technique called V arBaT ac [6]. Subsequently, a number of energyefficient communication schemes [7], [8] have been proposed by various authors. All such schemes follow some common approach in that the message to be transmitted is first encoded by a suitable source coding technique and after this encoding, one of the symbol values which will have the highest frequency of occurrence among the message symbols, will not be transmitted at all, i.e., the transmitter will be kept silent during the periods of this dominant symbol having the highest frequency of occurrence. Also, all these schemes utilize a hybrid Frequency-Shift Keying(F SK) - AmplitudeShift Keying (ASK) modulation/demodulation technique for the transceiver design. The RBNSiZeComm [7] scheme uses a source coding technique based on redundant binary number system (RBNS) so that the message is encoded using three symbol values - 0, 1 and $\bar{1}$. The transmitter is kept silent during the 0's only with a resulting energy saving of 53% at the transmitter over the conventional binary frequency shift keying (BF SK) technique. There was, however, no energy saving at the receiver in the RBNSiZeComm scheme. An altogether different encoding technique was used in the scheme Compression with Null Symbol (CNS) [8], where two consecutive message bits were encoded in terms of one of four possible symbol values, and one of these four symbol values was arbitrarily chosen as the silent symbol during which the transmitter was kept silent. This scheme gave rise to transmitter and receiver energy savings of 30% and 50%, respectively over BF SK. However, due to the requirement of increasing the battery life by a factor of ten in the upcoming 5G environment which is planned to be realized around 2020 with about 30 billion devices to be connected worldwide, more improved techniques for energy-efficient communication are 978-1-5386-6358-5/18/$31.00 ©2018 IEEE still called for to achieve further energy savings at transmitter and/or receiver.

1.1 Relevance of the Project

In order to monitor various environment conditions the need to deploy sensors has increased.

The data in the sensor cannot be recoverd once its battery life has expired, As the only solution is to replace the sensor in the wsn.

The redundant transmissions have to be decreased and made as minimal as possible.

1.2 Problem Statement

This project aims to provide a more efficient way to transfer data through wireless sensor networks.

We attempt to implement some of the algorithms and document the errors results and the conclusions we have observed.

1.3 Objectives

Objective is to reduce redundant transmission of data through wsn.

To implement different algorithms on our own and observe the results obtained and the challenges faced by using these algorithms.

To improve battery life of a sensor by grouping data

1.4 Scope of the Project

As the sensor nodes are mostly battery operated, design of energy efficient algorithms for routing in wireless sensor networks is a rapidly growing area of research. Presented research work focuses on the design of energy efficient algorithms for generic applications of WSN. However, the development of more energy efficient algorithms will result in increased application areas of WSNs in real time. In addition to the focus on the reduction in high frequency of reclustering and distribution of cluster heads, the areas such as finding the optimal route with high residual energy for inter-cluster multi-hop communication, placing multiple base stations to reduce the load on cluster heads close to a single base station may be considered to bring further improvement in energy efficiency and overall lifetime of the network in WSN.

1.5 Project Methodology

Agile:

Agile is a process by which a team can manage a project by breaking it up into several stages and involving constant collaboration with stakeholders and continuous improvement and iteration at every stage. It promotes **continuous iteration** of development and testing throughout the software development life cycle of the project. Both development and testing activities are concurrent.

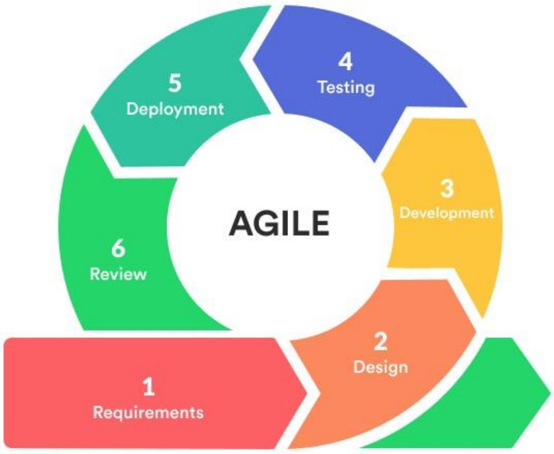


Fig 1.1 - Agile Methodology

Scrum

SCRUM is an agile development method which concentrates specifically on how to manage tasks within a team-based development environment. Scrum encourages teams to learn through experiences, self-organize while working on a problem.

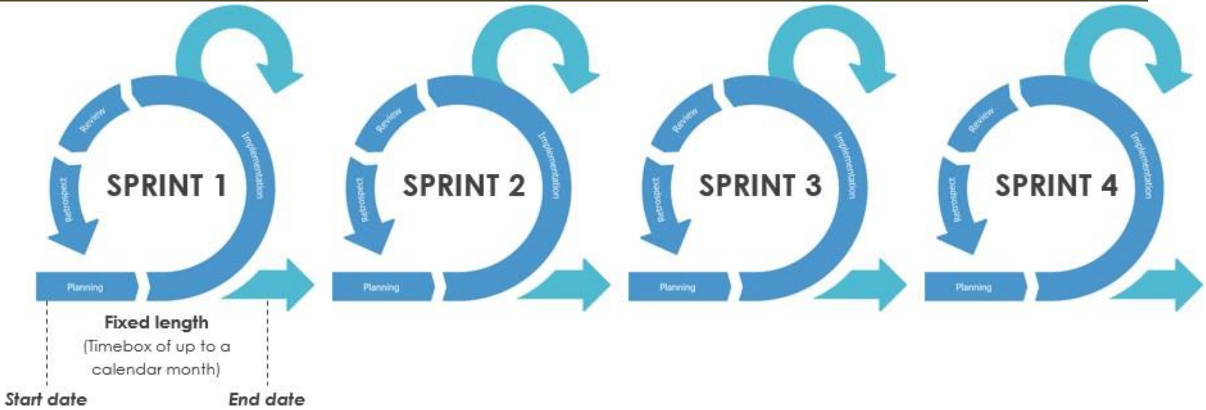


Fig 1.2 – Scrum

The Main Artefacts

- **Product Backlog** is the master list of work that needs to get done maintained by the product owner or product manager.
- **Sprint Backlog** is the list of items, user stories, or bug fixes, selected by the development team for implementation in the current sprint cycle.
- **Increment** (or Sprint Goal) is the usable end-product from a sprint.

Sprint 1	Sprint 2	Sprint 3	Sprint 4	Sprint 5	Sprint 6
August 2019	September 2019- October 2019	November 2019	December 2019	January 2020- February 2020	March 2020- April 2020
Discussion of project selection	Literature survey, Synopsis, SDD, Research Notes	Freezing of algorithm, ramework and software	Exams (hiatus)	Coding and implementation	Testing and adding last minute features

Table 1.1 – Sprint

CHAPTER 2

LITERATURE SURVEY

The following chapters give an overview of the various methodologies used by various authors for efficient energy transmission. We can observe that there is fine comparison made between major machine learning algorithms whether they are able to provide efficient energy emission techniques . The research efforts presented by the authors in the following papers are focused in developing and evaluating the data to do the same.

2.1 Mac protocols for wireless sensor networks: A survey

Approaches

Author: Demirkol, C. Ersoy, and F. Alagoz

Published In: IEEE Communications Magazine, vol. 44, no. 4, pp. 115–121, 2006.

The authors have used MAC Protocols for WSN [1]

Wireless sensor networks are appealing to researchers due to their wide range of application potential in areas such as target detection and tracking, environmental monitoring, industrial process monitoring, and tactical systems. However, low sensing ranges result in dense networks and thus it becomes necessary to achieve an efficient medium-access protocol subject to power constraints. Various medium-access control (MAC) protocols with different objectives have been proposed for wireless sensor networks. In this article, we first outline the sensor network properties that are crucial for the design of MAC layer protocols. Then, we describe several MAC protocols proposed for sensor networks, emphasizing their strengths and weaknesses. Finally, we point out open research issues with regard to MAC layer design.

2.2 Energy efficient collision aware multipath routing for wireless sensor networks

Authors: Z. Wang, E. Bulut, and B. K. Szymanski

Published In: Communications, 2009. ICC’09. IEEE International Conference on. IEEE, 2009, pp. 1–5

The authors design multipath routing for WSN [2]

Multipath routing can reduce the need for route updates, balance the traffic load and increase the data transfer rate in a wireless sensor network, improving the utilization of the limited energy of sensor nodes. However, previous multiple path routing methods use flooding for route discovery and transmit data with maximum power regardless of need, which results in waste of energy. Moreover, often a serious problem of collisions among multiple paths arises. In this paper, we propose an energy efficient and collision aware (EECA) node-disjoint multipath routing algorithm for wireless sensor networks. With the aid of node position information, the EECA algorithm attempts to find two collision-free routes using constrained and power adjusted flooding and then transmits the data with minimum power needed through power control component of the protocol. Our preliminary simulation results show that ECCA algorithm results in good overall performance, saving energy and transferring data efficiently

2.3 “Energy-efficient broadcasting in ˇ all-wireless networks

Authors: M. Cagalj, J.-P. Hubaux, and C. C. Enz

Published In: Wireless Networks, vol. 11, no. 1-2, pp. 177–188, 2005.

The author talks about broadcasting energy efficiently in WSN[3]

In all-wireless networks, minimizing energy consumption is crucial as in most cases the nodes are battery-operated. We focus on the problem of power-optimal broadcast, for which it is well known that the broadcast nature of radio transmissions can be exploited to optimize energy consumption. This problem appears to be difficult to solve [30]. We provide a formal proof of NP-completeness for the general case and give an NP-completeness result for the geometric case; in the former, the network topology is represented by a generic graph with arbitrary weights, whereas in the latter a Euclidean distance is considered. For the general case, we show that it cannot be approximated better than O(log N), where N is the total number of nodes. We then describe an approximation algorithm that achieves the O(log N) approximation ratio. We also describe a new heuristic, Embedded Wireless Multicast Advantage. We show that it compares well with other proposals and we explain how it can be distributed.

2.4 Energy-efficient route selection strategies for wireless sensor networks

Authors: D. J. Vergados, N. A. Pantazis, and D. D. Vergados

Published In: ” Mobile Networks and Applications, vol. 13, no. 3-4, pp. 285–296, 2008.

The authors propose various route selection strategies for WSN[4].

In wireless sensor networks WSNs, sensor devices have limited supply of energy. The sensor death due to dissipating battery energy is one of the fundamental design issues in WSNs. Hence, energy efficiency is argued to be the most important requirement for any protocol designed for WSNs. With sensors acting as routers to transport the packets from a source to a destination sensor, multipath protocols are used to discover multiple paths with the objective to improve the reliability, efficiency, and security in WSNs. Selecting the path that minimizes the rate of sensors death and extends the lifetime of the network is the main challenge for multipath protocols. In this paper, we propose a new energy and security aware route selection ESARS protocol for WSNs. The first part of ESARS selects a route that maximizes the network lifetime based on a novel metric. The second part of ESARS finds the optimal security level for the selected path based on the estimated security risk of the path. Traditionally, these two parts are addressed separately in the literature, and this paper combines the two parts in one protocol. The proposed protocol is evaluated and compared with other protocols using both analytical analysis and extensive simulations. The results show that the proposed protocol not only achieves its main objective to extend the network lifetime by significantly reducing the sensors death rate but also uses the most optimal security level for the selected route. Moreover, in ESARS protocol, several threshold parameters were employed to provide flexibility per the needs of the application in which the sensors are used.

2.5 Challenges: communication through silence in wireless sensor networks

Authors: Y. Zhu and R. Sivakumar

Published In: Proceedings of the 11th annual international conference on Mobile computing and networking. ACM, 2005, pp. 140– 147.

The authors propose challenges in WSN[5]

Wireless sensor networks (WSNs) are typically characterized by a limited energy supply at sensor nodes. Hence, energy efficiency is an important issue in the system design and operation of WSNs. In this paper, we introduce a novel communication paradigm that enables energy-efficient information delivery in wireless sensor networks. Compared with traditional communication strategies, the proposed scheme explores a new dimension - *time*, to deliver information efficiently. We refer to the strategy as *Communication through Silence* (CtS). We identify a key drawback of CtS - *energy - throughput trade-off*, and explore optimization mechanisms that can alleviate the trade-off. We then present several challenges that need to be overcome, primarily at the medium access control layer of the network protocol stack, in order to realize CtS effectively.

CHAPTER 3

SYSTEM REQUIREMENTS SPECIFICATION

3.1 Functional Requirements – Dataset preparation and pre processing

3.1.1 Data Collection

Data collection is defined as the procedure of collecting, measuring and analysing accurate insights for research using standard validated techniques. A researcher can evaluate their hypothesis based on collected data. In most cases, data collection is the primary and most important step for research, irrespective of the field of research. The approach of data collection is different for different fields of study, depending on the required information. The most critical objective of data collection is ensuring that information-rich and reliable data is collected for statistical analysis so that data-driven decisions can be made for research.

3.1.2 Data Visualization

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. Example -

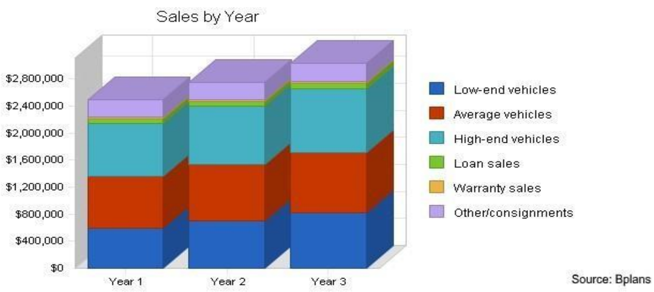


Fig 3.1 – Data Visualization

3.1.3 Data Labelling

Supervised machine learning, which we'll talk about below, entails training a predictive model on historical data with predefined target answers. An algorithm must be shown which target answers or attributes to look for. Mapping these target attributes in a dataset is called labelling. Data labelling takes much time and effort as datasets sufficient for machine learning may require thousands of records to be labelled. For instance, if your image recognition algorithm must classify types of bicycles, these types should be clearly defined and labelled in a dataset.

3.1.4 Data Selection

Data selection is defined as the process of determining the appropriate data type and source, as well as suitable instruments to collect data. Data selection precedes the actual practice of data collection. This definition distinguishes data selection from selective data reporting (selectively excluding data that is not supportive of a research hypothesis) and interactive/active data selection (using collected data for monitoring activities/events, or conducting secondary data analyses). The process of selecting suitable data for a research project can impact data integrity. After having collected all information, a *data analyst* chooses a subgroup of data to solve the defined problem.

For instance, if you save your customers' geographical location, you don't need to add their cell phones and bank card numbers to a dataset. But purchase history would be necessary. The selected data includes attributes that need to be considered when building a predictive model.

3.1.5 Data Pre-processing

Data pre-processing is a data mining technique that involves transforming raw data into an understandable format. Real-world data is often incomplete, inconsistent, and/or lacking in certain behaviours or trends, and is likely to contain many errors. Data preprocessing is a proven method of resolving such issues. The purpose of pre-processing is to convert raw data into a form that fits machine learning. Structured and clean data allows a data scientist to get more precise results from an applied machine learning model. The technique includes data formatting, cleaning, and sampling.

Data formatting. The importance of data formatting grows when data is acquired from various sources by different people. The first task for a data scientist is to standardize record formats. A specialist checks whether variables representing each attribute are recorded in the same way. Titles of products and services, prices, date formats, and addresses are examples of variables. The principle of data consistency also applies to attributes represented by numeric ranges.

Data cleaning. This set of procedures allows for removing noise and fixing inconsistencies in data. A data scientist can fill in missing data using imputation techniques, e.g. substituting missing values with mean attributes. A specialist also detects outliers — observations that deviate significantly from the rest of distribution. If an outlier indicates erroneous data, a data scientist deletes or corrects them if possible. This stage also includes removing incomplete and useless data objects.

Data anonymization. Sometimes a data scientist must anonymize or exclude attributes representing sensitive information (i.e. when working with healthcare and banking data).

Data sampling. Big datasets require more time and computational power for analysis. If a dataset is too large, applying data sampling is the way to go. A data scientist uses this technique to select a smaller but representative data sample to build and run models much faster, and at the same time to produce accurate outcomes.

3.1.6 Data Transformation

Data transformation is the process of converting data from one format or structure into another format or structure. Data transformation is critical to activities such as data integration and data management. Data transformation can include a range of activities: you might convert data types, cleanse data by removing nulls or duplicate data, enrich the data, or perform aggregations, depending on the needs of your project.

Scaling. Data may have numeric attributes (features) that span different ranges, for example, millimetres, meters, and kilometres. Scaling is about converting these attributes so that they will have the same scale, such as between 0 and 1, or 1 and 10 for the smallest and biggest value for an attribute.

Decomposition. Sometimes finding patterns in data with features representing complex concepts is more difficult. Decomposition technique can be applied in this case. During decomposition, a specialist converts higher level features into lower level ones. In other words, new features based on the existing ones are being added. Decomposition is mostly used in time series analysis. For example, to estimate a demand for air conditioners per month, a market research analyst converts data representing demand per quarters.

Aggregation. Unlike decomposition, aggregation aims at combining several features into a feature that represents them all. For example, you have collected basic information about your customers and particularly their age. To develop a demographic segmentation strategy, you need to distribute them into age categories, such as 16-20, 21-30, 31-40, etc. You use aggregation to create large-scale features based on smallscale ones. This technique allows you to reduce the size of a dataset without the loss of information.

3.1.7 Data Splitting

A dataset used for machine learning should be partitioned into three subsets — training, test, and validation sets.

Test set. A test set is needed for an evaluation of the trained model and its capability for generalization. The latter means a model's ability to identify patterns in new unseen data after having been trained over a training data. It is crucial to use different subsets for training and testing to avoid model overfitting, which is the incapacity for generalization we mentioned above.

Validation set. The purpose of a validation set is to tweak a model's hyperparameters — higher-level structural settings that cannot be directly learned from data. These settings can express, for instance, how complex a model is and how fast it finds patterns in data.

The proportion of a training and a test set is usually 80 to 20 percent, respectively. A training set is then split again, and its 20 percent will be used to form a validation set. At the same time, machine learning practitioner Jason Brownlee suggests using 66 percent of data for training and 33 percent for testing. A size of each subset depends on the total dataset size

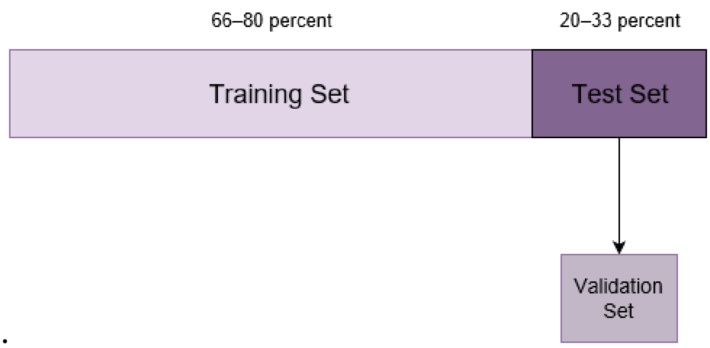


Fig 3.2 – Data Splitting

The more training data a data scientist uses, the better the potential model will perform. Consequently, more results of model testing data lead to better model performance and generalization capability.

3.1.8 Modelling

After pre-processing the collected data and split it into three subsets, we can proceed with a model training. This process entails "feeding" the algorithm with training data. An algorithm will process data and output a model that is able to find a target value (attribute) in new data — an answer you want to get with predictive analysis. The purpose of model training is to develop a model.

Two model training styles are most common — supervised and unsupervised learning. The choice of each style depends on whether you must forecast specific attributes or group data objects by similarities.

Model evaluation and testing: The goal of this step is to develop the simplest model able to formulate a target value fast and well enough. A data scientist can achieve this goal through model tuning. That is the optimization of model parameters to achieve an algorithm's best performance. One of the more efficient methods for model evaluation and tuning is cross-validation.

Cross-validation: Cross-validation is the most used tuning method. It entails splitting a training dataset into ten equal parts (folds). A given model is trained on only nine folds and then tested on the tenth one (the one previously left out). Training continues until every fold is left aside and used for testing. As a result of model performance measure, a specialist calculates a cross-validated score for each set of hyperparameters. A data scientist trains models with different sets of hyperparameters to define which model has the highest prediction accuracy. The cross-validated score indicates average model performance across ten hold-out folds.

3.1.9 Model Deployment

Deployment is the method by which you integrate a machine learning model into an existing production environment to make practical business decisions based on data. It is one of the last stages in the machine learning life cycle and can be one of the most cumbersome. Often, an organization's IT systems are incompatible with traditional model-building languages, forcing data scientists and programmers to spend valuable time and brainpower rewriting them.

3.2 Non-Functional Requirements

3.2.1 Usability

The system should be easy to use. The system also should be user friendly for users because anyone can use it instead of programmers.

3.2.2 Reliability

This software will be developed with machine learning, feature engineering and deep learning techniques. So, in this step there is no certain reliable percentage that is measurable. Also, user provided data will be used to compare with results and measure reliability. With recent machine learning techniques, user gained data should be enough for reliability if enough data is obtained.

3.2.3 Performance

Processing time and response time should be as little as possible providing the result at a faster rate when compared to other methods.

3.2.4 Supportability

The system should require Python knowledge to maintenance. If any problem acquire in user side and deep learning methods, it requires code knowledge and deep learning background to solve.

3.3 Hardware Requirements

- OS: Windows 10
- RAM: Minimum of 4GB

3.4 Software Requirements

- Basic Text-Editor: Visual Studio Code, Pycharm.
- Jupyter Notebook: Development of Python Scripts.

3.4.1 Why Python?

- General purpose programming language
- Increasing popularity for use in data science
- Easy to build end-to-end products like web applications
- Since the goal of this project is to build a web application, Python is a better choice. Though frameworks like Shiny can be used with R to create web applications, it is extremely slow.

There are two major libraries for machine learning in python: TensorFlow, ScikitLearn. The main differences are discussed below.

TensorFlow	Scikit-Learn
Low-Level, algorithms should be implemented manually	Provides off-the-shelf algorithms
Better choice for deep learning	Less efficient for deep learning
Steep learning curve	Easier to learn

Table 3.1 – Machine Learning Libraries

Scikit-learn was chosen for the project for the following reasons:

- Off-the-shelf algorithms
- Shallow learning curve compared to TensorFlow

CHAPTER 4

SYSTEM ANALYSIS AND DESIGN

Consider two binary strings S1 and S2, each of length n, to be transmitted to a receiver. As an example, these strings S1 and S2 may be received by an intermediate node of WSN from two sensor nodes in the network, and this intermediate node needs to transmit both S1 and S2 to another node. If each of these strings S1 and S2 is transmitted separately by the intermediate node, then the transmission of both S1 and S2 will require 2n time periods (one period refers to the time for transmitting one bit) to compute. However, in the proposed scheme, the message bits of both S1 and S2 will be transmitted in n time periods by a suitable encoding of the bits of S1 and S2 and at the same time, we would adopt the concept of silent communication as given in [8] to reduce the overall transmitter energy. Thus, our proposed scheme aims at reducing both the transmission energy as well as time for transmitting the bits of both the strings S1 and S2 taken together. To implement the above idea, we generate a composite string S12 from S1 and S2 by interleaving the bits of S1 and S2. Thus, if S1 = S1,n−1S1,n−2 · · · S1,1S1,0 and S2 = S2,n−1S2,n−2 · · · S2,1S2,0, then the composite string will be S12 = S1,n−1S2,n−1S1,n−2S2,n−2 · · · S1,1S2,1S1,0S2,0. In the next step, we encode each two consecutive bits S1,iS2,i (0 ≤ i ≤ n − 1) by one of the four symbols A, B, C and D according to the following rule: Encoded symbol is A or B or C or D when S1,iS2,i is equal to 00 or 01 or 10 or 11, respectively. By this process, we generate a block-encoded string S b 12 of n symbols from S12 where each symbol of S b 12 can take one of the four possible values. In a given block-encoded composite string, we find out the symbol having the maximum frequency of occurrence and decide to keep the transmitter silent during the periods of all occurrences of this symbols of higher frequency, following the concept of silent communication as used in [8]. We define this symbol for which the transmitter will be kept silent as the null symbol. The null symbol for each string will be dynamically chosen to get the maximum benefit of the silent communication in transmitter energy with a very little overhead in the transmitter and receiver design. For equal likelihood of all possible values of the bit strings S1 and S2, we estimate through experiments that the average value of maximum frequency of occurrence of a symbol in a composite block-encoded string S b 12 is about 40% when n grows very large (table.I). Thus, for large values of n, the transmitter will be silent on an average for 40% of the symbol periods. Along with this, transmission time will be reduced from 2n to n symbol periods

CHAPTER 5

PROBLEM FORMULATION

This paper presents an energy-efficient communication scheme called Dual-message Compression with Variable Null Symbol (DCV NS) using a novel source-coding technique to jointly encode two messages each of length n, which are to be sent to some receiver node. The source coding technique first interleaves the bits from the two messages and then encodes every two consecutive bits of this interleaved dualmessage by one of four symbol values so that the overall message length is reduced by a factor of two. The transmitter is kept silent during the periods of the most dominant symbol of the encoded dual-message. A simple hybrid modulation/ demodulation technique using non-coherent F SK and ASK is used for transmission and reception to achieve transmitter energy savings over that with conventional Binary FrequencyShift Keying (BF SK) of about 43.46% for equal likelihood of all possible messages and 54.1% to 58.9% for typical reallife sensor data, over a channel with Additive White Gaussian Noise (AW GN). Also, at the receiver side, there will be 50% energy savings over BF SK. The proposed approach is implementable at low cost and with low circuit complexity due to the use of the non-coherent BF SK receiver which would make it very suitable for those applications where a receiver node receives data from two sensor nodes which are further sent to a sink node through multihop communication. Further, due to the reduction in message length by a factor of two, this scheme also enables us to reduce the latency in communicating the data packets in such applications.

NAME	SUMMARY
ILETS	Collection of latitudes and longitudes.
Temputarue (ilets)	Collectionof temperature of wsn.
Time (ilets)	This dataset comprises of the time.
Latitude (ilets)	Contains latitude information of wsn.
Longitude (ilets)	Contains longitude information of wsn.

CHAPTER 6

IMPLEMENTATION

6.1 Clustering

Cluster analysis or clustering is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense) to each other than to those in other groups (clusters).

Cluster analysis itself is not one specific algorithm, but the general task to be solved. It can be achieved by various algorithms that differ significantly in their understanding of what constitutes a cluster and how to efficiently find them.

Basically there are 3 types of clusters, Fail-over, Load-balancing and HIGH Performance Computing, The most deployed ones are probably the Failover cluster and the Load-balancing Cluster. Fail-over Clusters consist of 2 or more network connected computers with a separate heartbeat connection between the 2 hosts

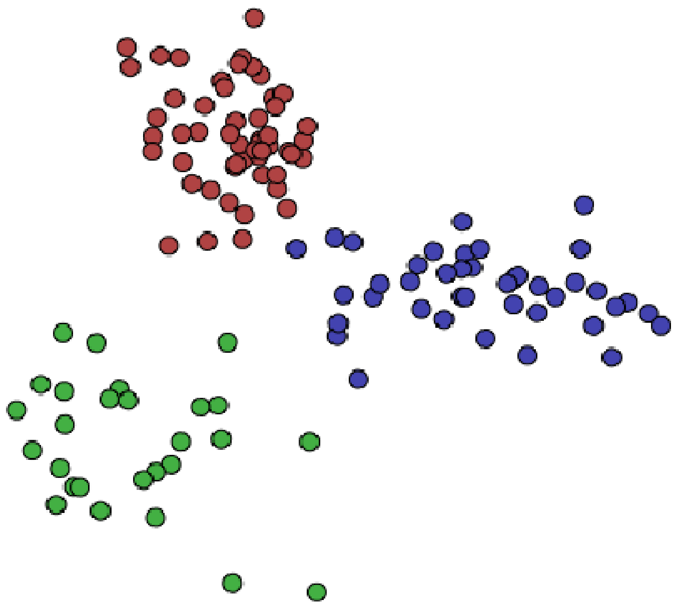


Fig 6.1 Clustering

6.2 K Means

The K-means algorithm starts by randomly choosing a centroid value for each cluster. After that the algorithm iteratively performs three steps:

Find the Euclidean distance between each data instance and centroids of all the clusters.

Assign the data instances to the cluster of the centroid with nearest distance.

Calculate new centroid values based on the mean values of the coordinates of all the data instances from the corresponding cluster.

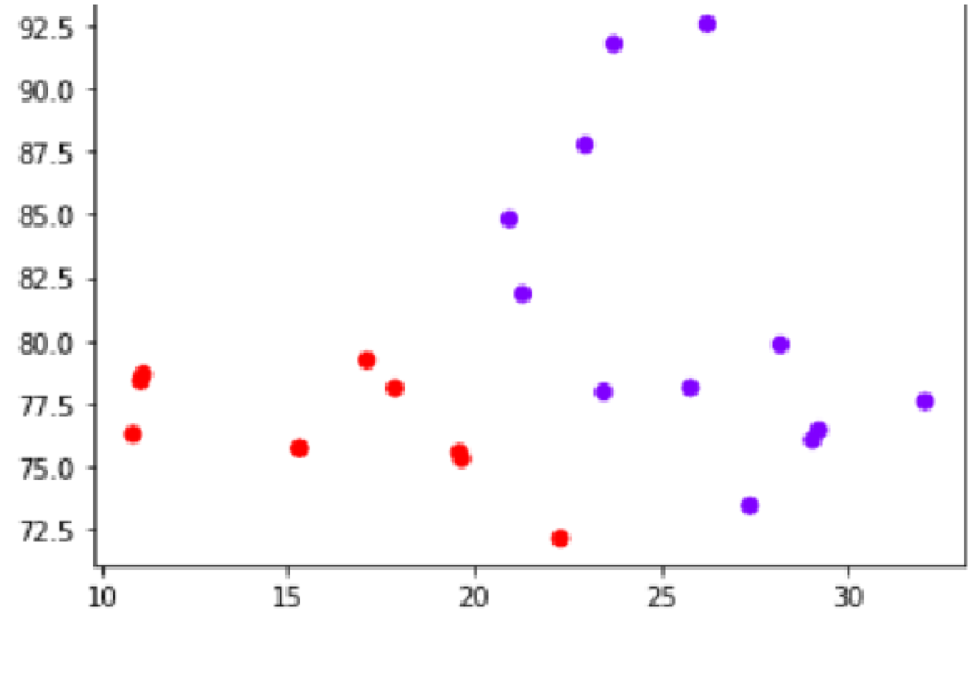


Fig 6.2 K Means

6.3 Euclidean Distance

Compute the Euclidean distance for one dimension. The distance between two points in one dimension is simply the absolute value of the difference between their coordinates. Mathematically, this is shown as |p1 - q1| where p1 is the first coordinate of the first point and q1 is the first coordinate of the second point.

Euclidean Distance from Cluster Centroid c1 = (16.8,17)	Euclidean Distance from Cluster Centroid c2 = (70.2,74.2)
18.30 7.08 5.31	96.54 84.43 83.16
10.04	79.09
30.95	49.68
86.37	15.38
83.10	5.85

Fig 6.3. Euclidean Distance

6.4 Proposed Physical System

To implement the above idea, we consider below the design of the proposed transceiver and also the hardware circuit for dynamic selection of the null symbol for a given pair of strings S1 and S2.

A. Transceiver Design As the transmitter will not transmit during the periods of the symbol values having maximum frequency of occurrence in the string S b 12, the modulation scheme to be used must have the flavor of ASK which has a high bit error rate (BER). To improve this BER, we propose to use FSK in conjunction with ASK leading to a hybrid FSK-ASK modulation scheme. Thus, we would use three carrier frequencies fc1, fc2 and fc3 to transmit three out of the four symbol values in

S b 12, and the transmitter will be kept silent during the periods of the remaining symbol value having maximum frequency of occurrence. A representative schematic diagram of this hybrid FSK-ASK receiver is shown in Fig. 1, in which r is the signal received at the input of the receiver which consists of the transmitted signal s and the noise η, i.e., r = s + η, and the outputs Ss1, Ss2 and Ss3 correspond to the symbols which were transmitted with the frequencies fc1, fc2 and fc3, respectively. The output Snull corresponds to the null (silent) symbol chosen during transmission. The received signal r is first passed through three bandpass filters (BP F) with center frequencies fc1 , fc2 and fc3 , respectively. The outputs of the three bandpass filters are applied to the three envelope detectors (ED) to get the values es1 , es2 and es3 corresponding to the three ON state symbols. In the ideal case, i.e., when there will be no noise (η = 0), one of these es1 , es2 and es3 will assume a non-zero value depending on the ON state transmitted symbols, while all of es1 , es2 and es3 will be zero if no symbol was transmitted during a symbol period. To detect the received symbol correctly, first es1 , es2 and es3 will be passed through a comparator to detect the largest values among es1 , es2 and es3 . Also, these signals will be passed through three threshold detectors to check if any of these es1 , es2 and es3 are above a certain threshold value eth. The outputs of the three threshold detectors i.e., qs1 , qs2 and qs3 are ORed together to get the value q. If the transmitter is in silent state during transmission of silent symbol, then all the outputs of the threshold detectors will be zero, making q to be zero. If the transmitted symbol is any one of the ON state symbols, then only Ss1 (or Ss2 or Ss3 , respectively) will assume the value logic 1. Ss1 Ss2 Ss3 Snull r = s + η fc1 fc2 fc3 ED ED ED BPF eS1 eS2 eS3 eth eth eth Comparator q qS1 qS2 qS3 eˆS1 eˆS2 eˆS3 Fig. 1. Schematic diagram of non-coherent F SK − ASK receiver scheme

B. Performance Comparison in a Noisy Channel We now analyze the performance of our proposed scheme as follows. The channel noise is assumed to be Additive White Gaussian Noise (AW GN).

C. If k bits are blocked together to generate a symbol, then the bit error rate (BER) is related to the symbol error rate (SER) by the relation,

BER = 2 k−1 2 k − 1 .Ps

D. BER for Different SNR Values:

To compute the bit error rate for different SNR values from the above equation, we set ss1 = 1, i.e., signal power = s 2 s1 2 and SNR value = 1 2.σ2 . The computations to find the optimum value of threshold voltage eth that gives the minimum BER for a given SNR value were implemented using Matlab.

E. Effect of Null Symbols on Energy Savings:

Note that from eqn.15, we get the required peak transmitter power for a given BER value, which corresponds to the scenario where the transmitter is switched ON for the whole duration of message transmission.

F. Effect of Data Compression on Energy Savings:

We now consider the effect of data compression on the overall transmitter energy. Let Pb and PD be the transmitter energy requirements for binary FSK and DCV NS respectively.

Hence we have, ED = PDTD = 1.1309 2 .PbTb = 0.5654Eb

Fig 6.4 Dataset

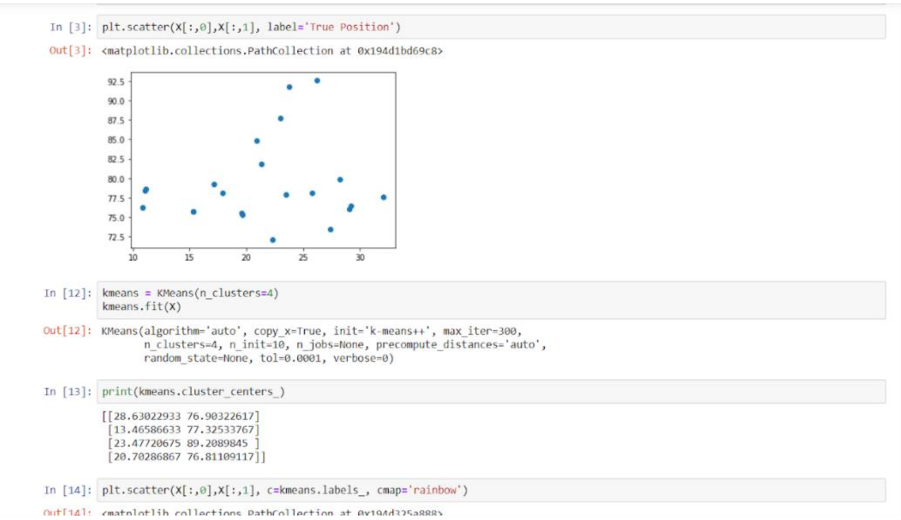
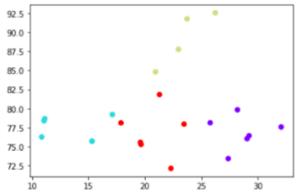


Fig 6.5 Scatter Plot formation

```
In [14]: plt.scatter(X[:,0],X[:,1], c=kmeans.labels_, cmap='rainbow')
Out[14]: <matplotlib.collections.PathCollection at 0x194d325a888>
```




```
In [17]: plt.scatter(X[:,0], X[:,1], c=kmeans.labels_, cmap='rainbow')
         plt.scatter(kmeans.cluster_centers_[:,0] ,kmeans.cluster_centers_[:,1], color='black')
Out[17]: <matplotlib.collections.PathCollection at 0x194d33a41c8>
```

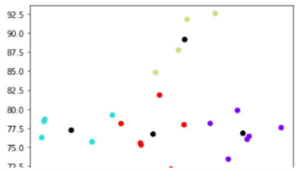


Fig 6.6 Formation of Clusters

CHAPTER 7

RESULTS AND DISCUSSION

7.1 Result

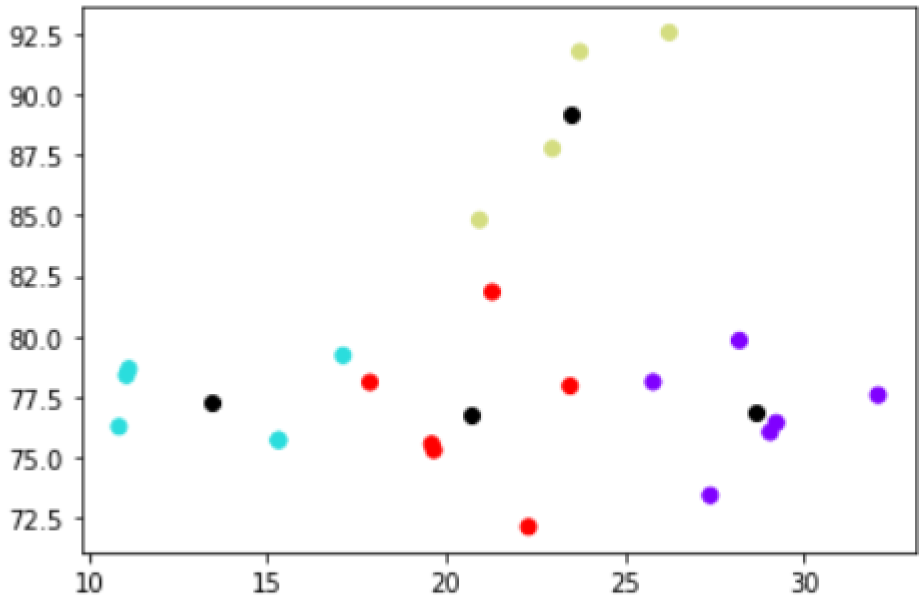


Fig 7.1 Formation of Cluster head

7.2 Discussion

It is noticed that 50% of redundant transmissions are avoided through this technique.

As the battery life of a wireless sensor cannot be extended it is very important to conserve energy during data transmission.

Clustering is a very efficient technique to breakdown large sets of data into smaller chunks and transmit then efficiently using cluster head.

CHAPTER 8

CONCLUSION

In this paper, we have proposed an energy-efficient communication scheme that combines the power of data compression based on encoding the messages by a novel source coding technique, called Dual-message Compression with Variable Null Symbol (DCV NS) and communication through silence. A simple hybrid modulation / demodulation technique using non-coherent F SK and ASK is used to have 43.46% energy savings on an average at the transmitter using a channel with Additive White Gaussian Noise (AW GN), and 50% energy savings at the receiver over the conventional Binary Frequency Shift Keying (BF SK) system, with equal likelihood of all possible strings of binary messages. However, simulation results with typical real-life sensor data show that, in practice, we can get about 54.1% to 58.9% energy savings at the transmitter over conventional BFSK. In contrast to the existing similar schemes, the null symbol is chosen dynamically to have the maximum advantage in saving transmitter energy. Also, the transmission time is reduced by a factor of two by reducing length of the encoded message by a factor of two, which results in reduced latency for message transmission. The proposed DCV NS scheme is very suitable for low cost low power WSNs using multihop communication

REFERENCES

[1] I. Demirkol, C. Ersoy, and F. Alagoz, "Mac protocols for wireless sensor networks: a survey," IEEE Communications Magazine, vol. 44, no. 4, pp. 115–121, 2006.

[2] Z. Wang, E. Bulut, and B. K. Szymanski, "Energy efficient collision aware multipath routing for wireless sensor networks," in Communications, 2009. ICC'09. IEEE International Conference on. IEEE, 2009, pp. 1–5.

[3] M. Cagalj, J.-P. Hubaux, and C. C. Enz, "Energy-efficient broadcasting in ˇ all-wireless networks," Wireless Networks, vol. 11, no. 1-2, pp. 177–188, 2005.

[4] D. J. Vergados, N. A. Pantazis, and D. D. Vergados, "Energy-efficient route selection strategies for wireless sensor networks," Mobile Networks and Applications, vol. 13, no. 3-4, pp. 285–296, 2008.

[5] Y. Zhu and R. Sivakumar, "Challenges: communication through silence in wireless sensor networks," in Proceedings of the 11th annual international conference on Mobile computing and networking. ACM, 2005, pp. 140– 147.