Energy Efficient Data Dissemination for Large-Scale Smart Farming Using Reinforcement Learning

Author

Muhammad Yasir Ali

Registration Number

2017-NUST-Ms-IT-00000204881

Supervisor

Dr. Asad Waqar Malik
Department of Computing

SCHOOL OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD, PAKISTAN

February 2021

Energy efficient data dissemination for large-scale smart farming using reinforcement learning

Author

Muhammad Yasir Ali

Registration Number

2017-NUST-Ms-IT-00000204881

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Information Technology

Thesis Supervisor:

Dr. Asad Waqar Malik

Thesis Supervisor's Signature: ____

DEPARTMENT OF COMPUTER SCIENCE

SCHOOL OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY, ISLAMABAD

FEBRUARY, 2021

Approval

It is certified that the contents and form of the thesis entitled **"Energy Efficient Data Dissemination for Large-Scale Smart Farming Using Reinforcement Learning"** submitted by Muhammad Yasir Ali have been found satisfactory for the requirement of the degree.

Advisor: **Dr. Asad Waqar Malik___**

Signature: _______________

Date: ____________________

Committee Member 1: **Dr. Anis Ur Rahman____**

Signature: _______________

Date: ____________________

Committee Member 2: **Dr. Muhammad Shahzad __**

Signature: _______________

Date: ____________________

Committee Member 3: **Dr. Muazzam Ali Khan Khattak**

Signature: _______________

Date: ____________________

Declaration

I certify that this research work titled "*Energy efficient data dissemination for large-scale smart farming using reinforcement learning"* is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources it has been properly acknowledged / referred.

Signature of Student

Muhammad Yasir Ali

2017-NUST-Ms-IT-00000204881

Plagiarism Certificate (Turnitin Report)

This thesis has been checked for Plagiarism. Turnitin report endorsed by Supervisor is attached.

Signature of Student

Muhammad Yasir Ali

Registration Number

2017-NUST-Ms-IT-00000204881

Signature of Supervisor

Dr. Asad Waqar Malik

Copyright Statement

Acknowledgements

I am thankful to my Creator Allah Subhana-Watala to have guided me throughout this work at every step and for every new thought which You setup in my mind to improve it. Indeed I could have done nothing without Your priceless help and guidance. Whosoever helped me throughout the course of my thesis, whether my parents or any other individual was Your will, so indeed none be worthy of praise but You.

I am profusely thankful to my beloved parents who raised me when I was not capable of walking and continued to support me throughout in every department of my life.

I would also like to express special thanks to my supervisor Dr. Asad Waqar Malik for his help throughout my thesis. I can safely say that it would not be possible without his guidance that he did during this pursuit.

Along with my advisor, I would like to acknowledge my entire thesis committee: Dr. Anis Ur Rahman, Dr. Muazzam Ali Khan Khattak and Dr. Muhammad Shahzad for their cooperation and prudent suggestions.

Finally, I would like to express my gratitude to all the individuals who have rendered valuable assistance to my study.

Dedicated to my exceptional parents: ***Mukhtar Hussain & Kubra Shaheen*** *and adored siblings whose tremendous support and cooperation led me to this wonderful accomplishment.*

Abstract

Smart farming is essential to increase the crop production for which IOT in agricultural parameters is crucial for the growth of crops. To achieve this, modern technology is required to enable accuracy in fertilizing, watering and adding pesticides to the crops as well as monitoring the conditions of environment. Now a days, more and more sophisticated sensors are developed but to achieve this on larger scale managing them efficiently is very significant.

We want to achieve sustainability in large scale farms by improving communication between the Wireless Sensor Network nodes and the Base Station through monitoring the energy and communication of sensor nodes through machine learning algorithms. The idea is to make multihop communication efficient, by balancing the energy consumption of sensors. The nodes with more energy will be used as relay to transmit data over the distance. The path selection will done based on remaining energy of the sensor nodes. Reinforcement learning is proposed to select best paths among the fields towards the base station. Reinforcement Learning is the area of machine learning in which is concerned with how involved agents are supposed to take actions in a specified environment to maximize the reward and to achieve a common goal. Reinforcement learning.

In our network, a large number of sensors are deployed on large scale fields, reinforcement learning is used to find the best path towards a base station. After a number of successful paths have been developed, they are then used to transmit the sensed data from the fields. The simulation results have shown better performance over shortest paths and broadcasting techniques that were tested against reinforcement learning.

Key Words: *wireless sensor network, reinforcement learning, simulation, AnyLogic, agent based modelling*

Table of Contents

List of Figures

List of Tables

Chapter 1: INTRODUCTION

1.1 Description and Motivation

Recently, there has been a large development in wireless sensor network domain. Wireless sensor networks have been deployed in almost every field. Agricultural wireless sensor networks have also been developed and deployed since a long time but they were only restricted in the form of a simple network that connects different sensors for the communication of sensed data. With the advancement in machine learning, there is an extensive application but for agricultural networks, they still have not been paid much attention to. There is quite a research potential in agricultural wireless sensor networks from networking point of view. So, the main motivation was to develop an agricultural network that can establish an optimal performing network for data dissemination among the field on large scale.

1.2 Problem Statement:

With the technology advancements and low cost sensors, more and more farmers are considering wireless sensor networks for the management of their crops [1]. In the existing agricultural work, networking elements like delay and path optimization etc. have not been considered much in large scale farming. The networks were focused mainly on physical deployment of sensors without managed networking which is not suitable for large scale farming because of cost and administrative reasons. If the sensor network paths are not managed efficiently then disconnected networks are created which are not helpful for large scale farming. To introduce efficient data sending over optimal paths based on the energy of the sensor nodes. We add a reinforcement learning algorithm that once finds a set of suitable paths which will help transfer data to the base station in an energy efficient manner.

1.3 Application:

The main purpose of this system is meant for wireless sensor networks that have their nodes deployed statically especially i.e. agricultural sensor network but with some modifications this mechanism can be applied to any application that contains wireless sensor networks. For example, this system with some optimization of the algorithm to accommodate mobility and it can be applied to wireless sensor networks for vehicular network and UAVs.

This system can also be applied to the wireless sensor network deployed for security systems where security devices with little no or mobility communicate sensor data to the server. This system can also be applied for military battlegrounds where sensors are deployed for spying or stealth purposes. Wireless sensor network is also applied to animal tracking system where animals are constantly monitored by sensors [2]. Intelligent smart home systems have also been proposed and developed by the researchers which are meant to monitor the home environments as well as entertainment purposes [3, 4]. Lim et al. suggested a monitoring system that could monitor the power system of a home [5]. It can regulate voltage and current system of a home. In general this system is applicable to any application that needs communication between wireless sensor nodes deployed statically and they are possibilities for the change in network topology if not frequently.

1.4 Aims and Objectives:

To conduct this research we have objectives to achieve:

- First to review and compare the most recent advancements in wireless sensor networks meant for agriculture. The analysis will also consider the latest wireless sensor network developments.
- Based on that analysis we will design and build the agricultural wireless sensor network for scale large farming which can build optimal paths to communicate sensed from the fields.

1.5 Organization of Thesis

The work on this thesis is organized as follows:

Chapter 2 consists of the discussion about wireless sensor networks, agricultural networks, machine learning and its types and its application in in wireless sensor networks.

Chapter 3 the covers the detailed review of literature and discusses the advancements and significant work done in in agriculture and wireless sensor networks.

Chapter 4 presents the proposed methodology.

Chapter 5 discusses the designing and implementation of the proposed methodology as well as the results gain from different experiments done on the developed model.

Chapter 6 consists of the conclusions drawn from the experiments and discusses the potential scope as the future work.

Chapter 2: Literature Review

The major concepts that were used in the project are discussed in this chapter. This chapter discusses the relevant background of the studies that were involved during the research work. Revision of Wireless Sensor Networks and their use in agriculture, Machine Learning and other relevant concepts are considered to cover the theory and literary background. This chapter also holds a review of the recent research papers that were considered in the study.

2.1 Wireless Sensor Networks

Wireless Sensor Network is a lot more pursued domain in networks. A Wireless Sensor network is defined an infrastructure less and self-configuring network where devices, nodes or machines are connected through radio signals. A wireless sensor node is equipped with sensing and computing devices, radio signal transmitters, receivers and a battery. The sensor node is usually a low power computing device and has limited resources in terms of processing, storage and communication range. The wireless nodes are connected through radio technologies like through Bluetooth connection, Wifi, ZigBee, LoRa WAN and etc. Wifi has a capability to monitor through video [6].ZigBee supports several network topologies like ring, star and mesh etc. [7]. ZigBee also is quite efficient and capable technology to capture soil parameters in agricultural networks [8]. LoRa is used where sensors cannot transmit more than some bytes in one time and LoRa handles this limitation of other technology with larger capacities [9]. After the nodes have been deployed in specific or random topology they are responsible for self-organizing and finding a path.   The communication usually happens in the form of multi hop fashion. The sensor nodes try to collect information and forward it to the main station the base station. Wireless sensor networks are deployed in harsh conditions where normal infrastructure cannot be installed. Battery is crucial resource for sensor nodes [10]. The sensor nodes are meant to live there for their complete life until dead. So, longer network life is a key for evaluating the sensor network in an application- specific way [11]. Due to the major uses of wireless sensor networks it has become important to maximize the lifetime of wireless sensor networks to obtain more and more information as possible [12]. To wireless sensor network considering the limited battery they to work with.

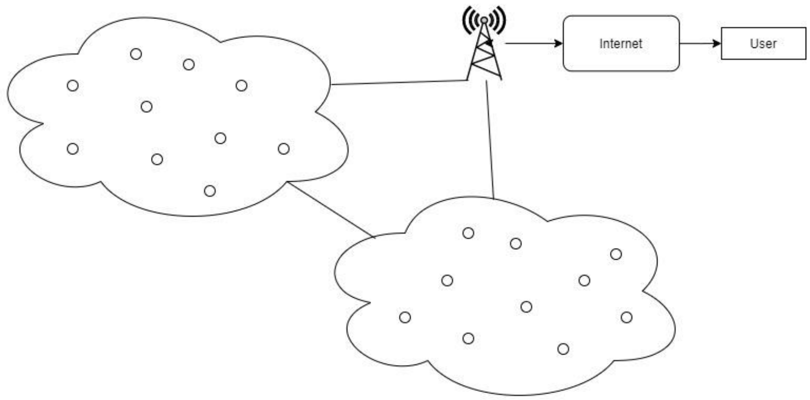


Fig. 1 Wireless Sensor Network [30]

The application of wireless sensor network is usually the tough environments where human reach is not easily possible. Real time monitoring and video surveillance over a long range has become possible due to wireless sensor networks [13]. The application includes smart buildings, military environments, precision agriculture, security and surveillance, smart grids, transportation and health care.

2.2 Machine Learning

In 1927, the British statistician Yule gave an Autoregressive model (AR model) to predict in law of market changes [14]. Walker established Moving Average model and later he combined the both models and established Autoregressive Moving Average (ARMA) model [15]. Box and Jenkins gave Autoregressive Integrated Moving Average Model (ARIMA) which was meant for time series prediction [16]. Combination with time series predictions in recent years, machine learning has become an active issue [17].

Machine Learning (ML) is rooted from Artificial Intelligence whose concept is first presented by Sir Arthur Samuel in 1959 [18]. The presented concept was similar to machine learning. According to the concept, Machine Learning allows a computer system to learn and decide without explicit programming. The concept was superficial and very fantastic at that time and a lot eager was seen to achieve this. In the current computing generation, it is being used to some extent as the concept stated at that time. This domain from Artificial Intelligence is different other computing fields because self-learning does not exist there while developed features are to be added in them. For example, in traditional software development process, changes in the code have to be made manually done by the programmer every time a new feature is to be introduced. Contrary to this, Machine Learning has another way of doing it because the set of instructions let the machine to learn new things by using the existing code on the machine. Many machine learning systems borrow their sequences form cognitive and thinking processes [19]. Machine Learning creates models are created based on the input data. The output is generated by those models in the form of decisions and predictions. In this way, a computer is able to make decisions whenever a new need or requirement arises. Machine Learning has three main categories. These categories are made based on how learning system executes the learning process. The categories include: supervised learning, unsupervised learning and reinforcement learning. An input with a labeled data is called supervised learning [20]. This type of learning model contains specific classes. Because it is supervised learning, the models adapts the inputs in such a way that it corresponds the outputs. When an unlabeled data is taken as input, it belongs to the unsupervised learning [21] [22]. The model learns the patterns from the input data. Reinforcement Learning is the third type of machine learning where an agent is given reward or punished based on decision it takes to find a specified goal [23] [24]. Here in this research we are pursuing the Reinforcement Learning because the model we are about to develop needs intelligent path finding.

2.3 Introduction of Machine Learning in Wireless Sensor Networks Meant for Agricultural Wireless Sensor Networks

The application of wireless sensor networks in IOT in agriculture helps the famers to know about their farms in quantitative and statistical manner which is quite useful in making the accurate management decisions [41]. As the number of interconnected devices grow then there is a need to manage big data and analyze it for useful decisions through spatial and temporal variations [42]. Machine learning has proven a lot more progress in almost every field and in wireless sensor networks as well. In this new era of technology machine learning can also be introduced in agriculture. For example machine learning can help us create a routing protocol which is energy efficient and time-saving for the whole infrastructure. This research focuses on the use of machine learning in in wireless sensor network which is actually meant for agriculture.

With the introduction of machine learning wireless sensor networks can be improved and the overall network efficiency can be enhanced. For example machine learning can add more active path finding and routing in between nodes. This can also help alleviate generic network problems like the hidden node problem. Machine learning can help find multiple routes among deployed nodes and if a node cannot access the base station directly then it can follow alternative paths that were developed by machine learning algorithm.

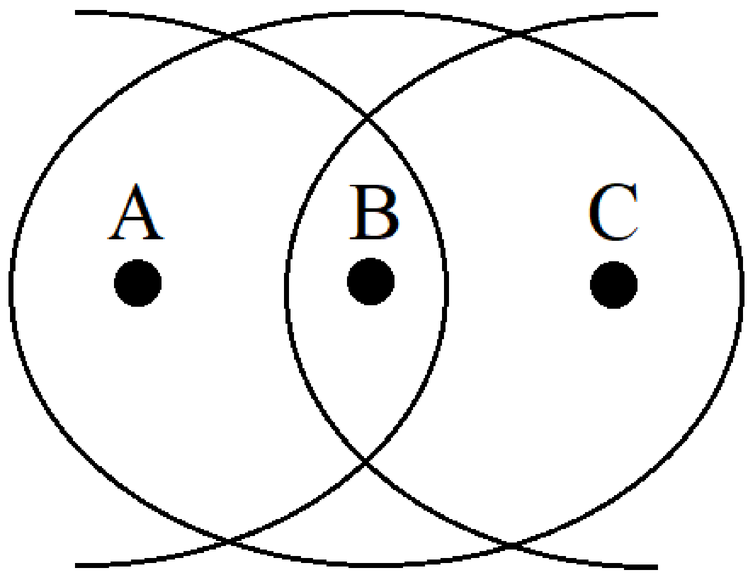


Fig. 2 the Hidden Node Problem [31]

Another problem is the energy hole problem where a set of intermediate nodes die and this creates several disconnected networks which are smaller in size and the communication between base station and those parted networks cannot happen because of low range of sensor nodes. As we know machine learning can help train the model through very diverse range of data sets and hidden node problem data set can also be trained to find an optimal set of paths for the wireless sensor network. To some extent machine learning can help alleviate this type of issue as well because machine learning algorithm can learn and create models in any given scenario, hence is solution can be devised for energy hole problem as well. The most important factor for wireless sensor network is the energy efficiency because the complete network relies on a limited resource of battery. There is already a lot of work in energy department for wireless sensor networks. Machine learning can also be applied to achieve efficiency energy in wireless sensor networks.

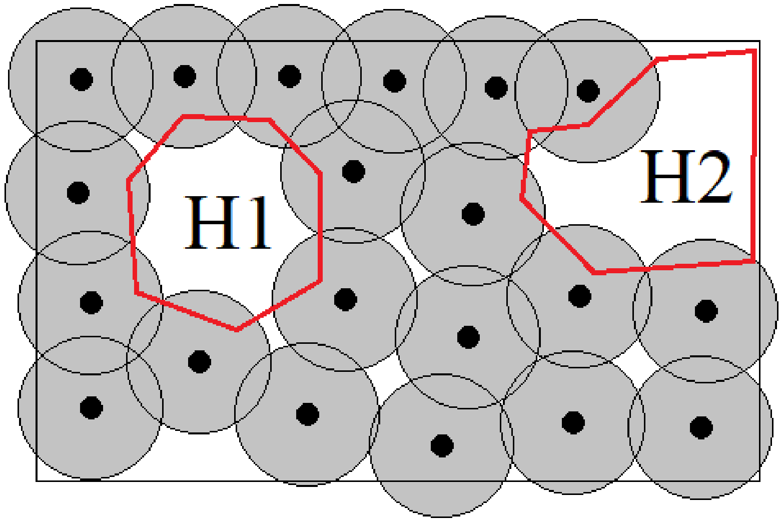


Fig. 3 Energy Hole Problem [32]

2.4 Theoretical Background of Wireless Sensor Networks

Here is brief background of Wireless Sensor Network that was meant for IOT in agriculture in different literature of the recent past.

Mare Srbinovska et. al. [33] proposed wireless sensor network, which is deployed to capture the information about environmental parameters that affect the crops. The objective was to develop a low power wireless sensor network that can live a longer life to collect information about the crops. They investigated the wireless sensor network and power consumption of nodes in detail to achieve longer network life and lower power consumption among the nodes that reside in the network. They aimed at developing a low cost non robust wireless sensor network that can be used for data collection of crops in green houses. They collected data from their deployed network and did a theoretical analysis of the power consumption from their network and compared with quantitative analysis from their deployed wireless sensor networks.

They did use the general working modes of wireless sensors. According to Mare Srbinovska et. al, the sensors should keep in in the sleeping mode as long as some external event triggers them to keep in in the processing mode where the sensors are about to sense collect and compute the data from the environment they are deployed in. The sensors should transmit this data to the server or the base station after processing the data. The energy consumption for communication or transmission mode is twice the energy consumption in processing mode. The relation for energy conservation is given below:

$$ETx = ETx\_s + ETx\_p + ETx\_RF$$

According to this relation the energy consumption of a sensor node (ETx) is the sum of energy consumption of each individual mode or component that is the processing module (ETx_p) and the radio frequency module (ETx_RF).

They deployed the wireless sensor network in pepper greenhouse to monitor the environmental effects [34]. The network was meant to monitor the parameters like temperature, PH, humidity and soil moisture. The whole architecture was meant to reduce the cost of environmental monitoring to manage the crops in a better way. Deployed a set of sensors and a base station, the data was sensed, collected and sent every 30 minutes, where the sensors were kept at the distance of 30 meters from each other. Sensors were kept in sleeping state and woke up every half of the hour to collect data. To achieve maximum life from the battery of the sensor they managed to keep sensors in sleeping mode as much as possible. To boost energy during day time, they made the use of solar cells with the sensors. The power consumption results from their deployed network were similar to the estimated theoretical analysis that they made.

P Jain et. al. [35] proposed energy efficient data collection from the wireless sensor network that involves a gateway in an agricultural environment. The sensor network optimizes itself to collect data from the nodes and it is in this way energy efficient. The gateway of the wireless sensor network adapts itself for they parameters that are needed to be sensed and the data is collected by sensor nodes. The adaptive learning of the gateway depends on the spatial and temporal crop characteristics that may change during the complete season. The gateway was introduced with supervised learning principles that were meant to set specific parameters for the sensor nodes. According to P Jain et. al. energy efficiency can be improved by specifically setting the sensing parameters at the gateway beforehand and communicate them with the sensor nodes so that they can sense and collect data about the very e specific parameters needed.

They analyzed the model which they compared with non-adaptive models as well. In non-adaptive model the data transmission occurred at continuous rate. The adaptive model had successfully shown the classification of data points for or diseased and non-diseased crops whose parameters were set by supervised learning at the gateway. With different variations in the data set the model was able to dynamically set the parameters for the crop. Results showed that the proposed adaptive sampling model based on Random Forest can significantly reduce energy consumption by varying the sampling rate of the sensor nodes. The adaptive model which was set on the basis of random forest algorithm gave 22% to 30% energy efficiency for the sampling rates kept more frequent than other techniques. Future work includes deployment of the adaptive model.

G Sathiya et. al [36] design a wireless sensor network with nodes and a base station using zigbee. The wireless sensor network had 7 nodes which comprised of six sensor nodes and one coordinator. The wireless sensor network was aimed at small-scale fields to sense and collect data about the crop. They successfully implemented the model collector results from the model. As their future work suggested, the nodes should be placed properly to achieve efficiency. Power consumption is also suggested as an extended work.

Zhang X et. al. [37] developed the wireless underground sensor network and implemented it in the university. Their network was fully operational and they were able to collect data on the website server. Their network model “Thoreau” was based on Sigfox Iot which operated at 902 MHz unlicensed ultra-narrow band. Their network had a backend system, a base station and a set of sensor nodes. The sensor nodes were based on the Texas Instrument MCU MSP430F5529 having the CC1120 radio frequency transceiver. Each sensor node had two sensor components (Decagon GS-3 and MPS-6) which were meant for collecting data for four crop properties: temperature, soil moisture, volumetric water content and electric conductivity. As a result, they collected so information and analyzed it at backend. The network was meant for small scale field and in future work they left it for larger scale work and it networking as well as energy efficiency parameters we are also to be considered.

Zainal A et. al. [38] studied and analyzed the performance of four routing algorithms meant for wireless sensor networks. The routing protocols they considered were Low Energy Adaptive Clustering Hierarchy (LEACH), Threshold sensitive Energy Efficient sensor Network (TEEN), Stable Election Protocol (SEP) and Energy Aware Multi Hop Multi Path (EAMMH) protocols. In Matlab, the simulation was done on the heterogeneous environment in which the performance of all these protocols were considered. The performance parameter that were considered for comparison were network lifetime, number of dead nodes per round, stability period of a network (the time till the death of first node), number of cluster heads per round, throughput and average residual energy per node. Their simulation results revealed that the stable region and network lifetime of TEEN is better than others while throughput from SEP was better than others.

Table 2.1. Literature review - Summary of Wireless Sensor Network incorporated in agriculture.

Title	**Objective**	**Limitation**	**Year of Publication**
Energy Consumption Estimation of Wireless Sensor Networks in Greenhouse Crop Production [33]	The objective was to develop a low power wireless sensor network that can live a longer life to collect information about the crops.	Network configuration, network lifespan, node optimization is neglected.	2017
Development of an Energy-efficient Adaptive IoT Gateway Model for Precision Agriculture [35]	Energy efficient data collection from the wireless sensor network that involves a gateway using supervised learning model	Network configuration, network lifespan, node optimization is neglected	2018
Designing a Wireless Sensor Network for Precision Agriculture Using Zigbee [36]	Aimed at small-scale fields to sense and collect data about the crop	Energy Efficiency, Network configuration, network lifespan, node optimization is neglected	2017
Thoreau: A subterranean wireless sensing network for agriculture and the environment [37]	Aimed at large-scale fields to sense and collect data about the crop	Energy Efficiency, Network configuration, network lifespan, node optimization is neglected	2017
Selection of energy efficient routing protocol for irrigation enabled by wireless sensor network [38]	Routing protocols comparison among LEACH, TEEN, SEP and EAMMH	Energy Efficiency not mentioned. Protocols were not meant for IOT in Agriculture.	2017

Chapter 3: Problem Statement and Proposed Model

Problem statement and problem formulation are explained in this chapter 3. Hair we will define the actual problem for which we have developed the solution. In in the later sections there is our proposed model.

3.1 Problem Statement

There has been a lot of work done in the domain of wireless sensor networks. Most of the wireless sensor networks are already developed and implemented in in in different fields like military, security, hospital and transport etc. but not a lot of attention has been paid to agriculture regarding wireless sensor networks. It can be denied that wireless sensor networks are deployed in agricultural domain but in their general configuration, that is without any proper optimisation for energy efficiency, network lifespan, sensor node optimisation and other networking elements like technology, architecture and introducing the latest technologies like machine learning into the agricultural domain for better management of crops. Secondly, the work that has already been done in regards to agricultural wireless sensor networking, the farms or fields taken into consideration were not of the large scale. As we have already seen from the literature review that some of the researchers have implemented their proposed network for agriculture but most of them were not considering large scale farming in terms of wireless sensor network. From the literature work, we have also seen that the networking architecture and topology was also not paid much attention to. Most of the work has only physical deployment of sensors without any managed networking which is not suitable for large scale farming because of cost and administrative reasons. If the sensor network paths are not managed efficiently then there is disconnected networks are created which is not helpful for large scale farming.

In order to achieve efficiency in terms of networking architectural elements and to alleviate energy consumption during path finding procedure, as well as to introduce new era technologies into the agricultural domain from networking perspective, we have to introduce efficient data sending over optimal paths based on the energy of the sensor nodes.

3.2 Reinforcement Learning

Reinforcement learning is the type of machine learning in which we take appropriate actions to increase the reward for a particular situation. Reinforcement learning is all about finding the best way or path that any individual may take in in some situation. Reinforcement learning is different from supervised learning because reinforcement learning does not have a specific set of data for training and individual has to take specific actions by itself for best behaviour whereas in supervised learning, there is already is dataset beforehand for training. In simple words reinforcement learning is the learning in which an individual learns the best behaviour by experience in any given situation.

3.2.1 Main Characteristics of Reinforcement Learning

- Input: Initial state is an input at the start of the model.
- Output: There are many possible outputs to a given situation.
- Training: the training is initiated from the input. The model gives back the experienced state for which there is a punishment on failed attempt or a reward on successful attempt towards the target in any given situation.
- There are a number of iterations till all possible outputs are found for any specific situation.
- The set of outputs having the maximum reward are considered as the best set of outputs.

3.2.2 Types of Reinforcement

There are two types of reinforcement.

Positive: positive reinforcement is an event that occurs because of a specific behavior and it increases the overall strength of the behavior. Positive reinforcement has a positive effect on the agent's behavior.

The advantages of positive reinforcement is that it increases the overall performance it can sustain changes for longer period of time the disadvantages include the overload of the states which can destroy the desired results.

Negative: negative reinforcement can be defined as the behavior that is strengthened due to some condition is avoided.

The negative reinforcement also increases the behavior and advantages include this setting of minimum standard to behave in a particular way. The disadvantages include very low criteria to meet minimum behavior.

3.2.3 Practical Uses of Reinforcement Learning

- Reinforcement learning is used in automation of robots in industries.
- Reinforcement learning is used for large scale data processing in machine learning applications.
- Reinforcement learning is used for the training of behaviors for providing the custom set of instructions and calculated amount of materials for a set of inputs in different industries.
- In education reinforcement learning is used to provide according to the requirement material and content to the students.

On larger scale reinforcement learning is used:

- In a model, where environment is known but the required solution is not known yet.
- Reinforcement learning is also used in environment where only simulation is given for any particular situation for which simulation-based optimization is needed.

3.3 Proposed Model

In order to achieve this solution to our proposed problem we have to create a simulation environment where we have mesh networks deployed in separate farms (roughly 370x330m), which are connected with a base station through wireless medium. The connections between farms and base station are single hop as well as multihop, based on their distance from base station. The nodes send data to the base station through optimal paths available based on their remaining energy.

The idea is to enable sustainability in large scale farming networks. In wireless sensor networks, we are focusing to increase lifetime to successfully send data to base station in an efficient manner.

There are several approaches to achieve multipath communication. We use machine learning or better stating, the reinforcement learning algorithm to create and select paths based on the remaining energy of the multihop sensor nodes. The energy information will be used to select maximum energy sensors to send the sensory data from the nodes deployed in the fields, towards the base station.

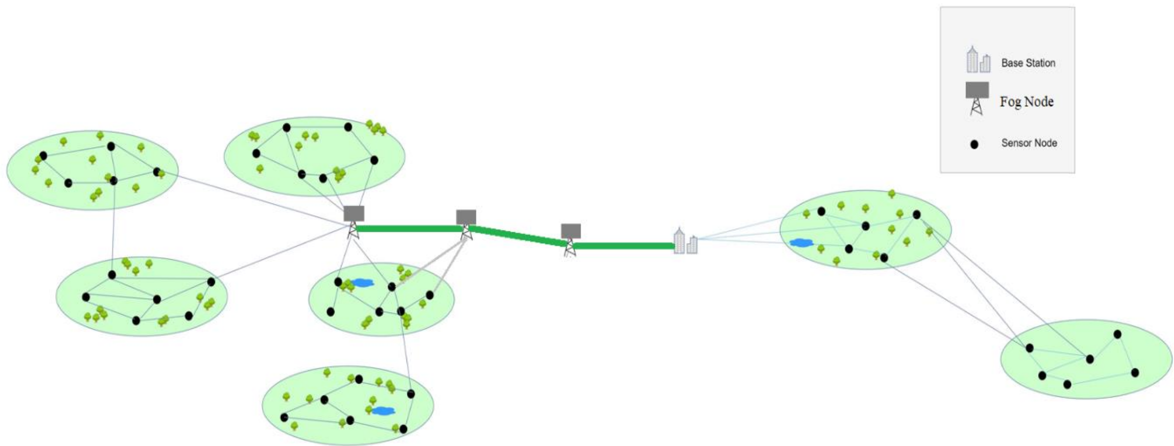


Fig. 4: Large Scale Farm Network Deployment

3.4 Summary

This chapter includes the proposed problem that we have defined to achieve network optimisation in terms of energy efficiency and optimise data sending towards the base station. Then we have shared a little knowledge about the technology e which will be used to develop this solution for the problem statement. After that there is a proposed model for that solution which consists of a number of fields having mesh networks. Those networks comprise of randomly deployed nodes on a larger scale. The path formulation for data sending is chosen through the reinforcement learning process. After the path formulation process has been done, those paths will be used to send the sensory data from nodes towards the base station periodically. The sensor cycles will be generic and the sensors will be sensing the data after specific period of time until the death.

Chapter 4: METHODOLOGY

This chapter presents the proposed methodology of the complete research work in detail. The details cover the design and implementation details of the project work done. The tools and techniques that were involved in making of the project work done.

4.1 Problem Statement

As we have seen in the previous chapters, for IOT in agriculture there is a need to introduce modern technologies for better crop management, to create sustainable and better use of resources. For this we need to upgrade agricultural management technologies. The literature review has shown that there is a lot of work have been done in this regard but actually without any proper management of networking and there is no optimization done in regards of network as well. This is because the major development was done by experts of electronics fields who focused mainly on the ground sensing architecture for agriculture let's say, the sensing infrastructure. The wireless sensor network have then been involved with their projects but in a very basic or generic forms. The actual data sending and sensor cycles are were needed to be optimized for better management and longer lifetime for the networks they used. Secondly, orientating the wireless sensor networks was also needed to be paid attention to. So, the problem statement of our research work states:

"To increase network lifetime, we want to achieve energy efficient data transmission in large scale farming using optimal paths developed through reinforcement learning."

Our main focus is to enable the sustainable networking which can handle the common problems that occur in networking of sensor nodes that are deployed on large scale and on random topology as well. The work that is reviewed in literature were mainly focusing the small scale of wireless sensor networks and only meant to provide basic functionality of data sending. For most reasons the whole architecture would also need an optimization. If we deploy the infrastructure meant for smaller scales, over a large spatial environments it will create complexities as it is not developed for that specific environment. So, in order to pay attention to large scale of IOT in agriculture there is already a good sensing infrastructure developed but in order to deploy that over a large scale was our main focus.

We did develop a simulation environment of the scenario that can help us achieve our goals. The development environment details will be discussed in the later sections.

4.2 Pre Processing of the Idea

We did a research project of modelling and simulation domain where we had to simulate an agricultural environment on large scale where a wireless sensor network is deployed. To consider the large scale scenario we did take a set of 6 fields where 370x330 meters is each field size. We needed to have a sensor deployment which are able to connect with each other wirelessly and can send sensory data towards the base station which is located at the end of these fields. The process of path finding between the nodes was done by reinforcement learning algorithm. The set of paths that were created as a result of reinforcement learning were used to send the data from the fields towards the base station.

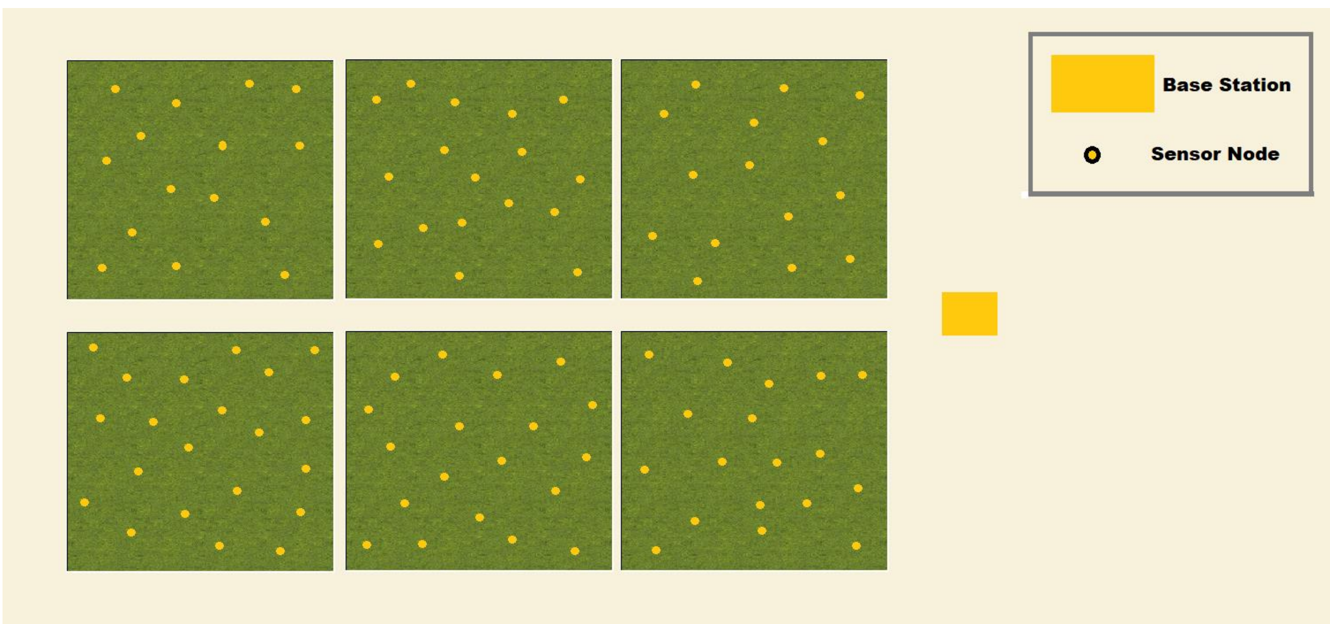


Fig. 5: Farming Scenario

4.3 Tools and Techniques

The selection of tools that are involved in the designing and implementation of the project are of crucial importance because there are certain elements connected to each other that may be impacted easily if parameters are changed during the research work. For our designing and implementation we did choose the following set of tools and software whose details are given below.

4.3.1 Computer Details:

We selected Hewlett-Packard EliteBook Folio 9480m with following configuration:

- **OS Name:** Microsoft Windows 10 Pro
- **Version:** 10.0.19041 Build 19041
- **System Manufacturer:** Hewlett-Packard
- **System Model:** HP EliteBook Folio 9480m
- **System Type:** x64-based PC
- **Processor:** Intel(R) Core(TM) i5-4310U CPU @ 2.00GHz, 2601 Mhz, 2 Core(s), 4 Logical Processor(s)
- **Total Physical Memory:** 7.42 GB
- **User Name:** DESKTOP-9JS3MO7\Si. Mirage

4.3.2 Software Details:

The software for development of the project was AnyLogic modelling and simulation software with following details.

AnyLogic is a powerful simulation and modelling software that is adopted in several industries worldwide for simulation solutions.

It is an agent based modelling tool which uses discrete event and system dynamics simulation techniques. AnyLogic is a cross platform tool and we did we use Windows version of this software.

AnyLogic supports many libraries to support multi environment simulations.

- AnyLogic 8 Personal Learning Edition 8.7.0
- Build: 8.7.0.202011181926 x64
- License Information

  Key type: File

  License Type: Personal Learning Edition

  Options: none

  Licensed to: Personal Learning Edition

  Company: AnyLogic North America

- Activation expires on: -
- Support Services expiration date: -

Visit https://www.anylogic.com

4.4 Design and Implementation

We did simulate a 6 farm field with a base station placed at the side of those fields. The number of short range sensor nodes deployed randomly and connected through wireless medium. We applied a Reinforcement learning algorithm before any sensing and sending of data can happen. The Reinforcement Learning algorithm helped us finding a suitable node which happened to be lying geographically nearest to the base station in each field. The rewards were given on each successful attempt to that node and the reward was considered as the successful path. As soon as the path had been developed, every sensor did start their sensor cycle to sense and send their data towards the base station using those paths.

The agent details that we did create for the project are given below.

- Main agent: Held all the basic simulation operations for the project.
- Message agent: This agent was meant to hold the messages which actually contained the

sensor id, sensor time consumption to send a specific message in time, energy consumption per message, as well as the field parameters like pH and temperature.

- Partition: This agent was actually the 370 x 330 meters field/ farm agent that created equal sized fields which contained the sensor agent as well.
- Pathfinder: This agent ran a reinforcement learning algorithm to find paths. This agent was meant to move from one sensor to another by checking their presence and availability and connecting them with the target node on successful attempts.
- Sensor: this agent it was meant for defining the properties for sensor nodes which would be residing inside the partition agent.
- Station: The station agent was meant for the base station purposes and all the messages were received at the base station where we did define database for saving all the data and exporting them to an excel file.

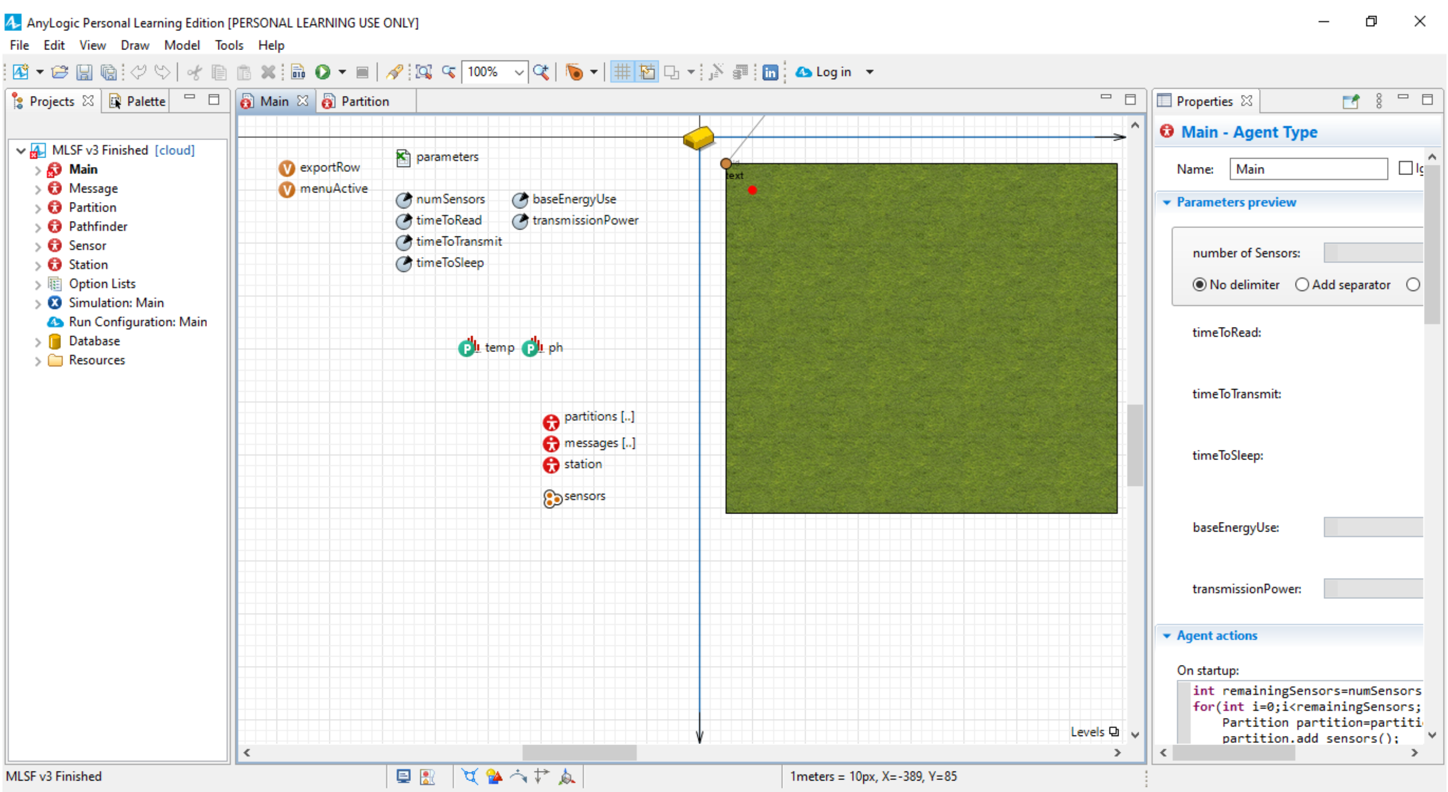


Fig. 6: Main agent image of the project under development.

4.4.1 Working of Reinforcement learning Algorithm

We had a main agent that had all the basic functionality defined for the model. There was also an agent named as sensor because that created the individual sensors in the eventual model. A sensor was an agent in and of itself. AnyLogic does actually offer to simulate in a cellular automaton but it's not applicable for this model because we wanted each sensor to know stuff about itself and that's not possible with a default setup so we went for a usual 2d environment and we had our sensor agents and we also did use a little line to show the learning that's going on by the help of the pathfinder agent. Each sensor was characterized by one parameter that was storing its type what type of sensor is it (there were two types of sensors either normal sensor which were visually shown with a green sphere or a gold colored sensor to show as an actual destination which was geographically located near the base station in each field), we did store this sensor type information using an option list in AnyLogic where basically we specified different options about the dimension we could have. The sensor also had two connection objects or link to agent objects, one was the connection to the pathfinder, which was only active when connecting to a pathfinder. The link to agents connection worked only when pathfinder was actually on the specific sensor at the moment otherwise there's no connection happening and then there was a persistent links to agents that stored the neighboring sensors.

The sensor agent stored an important data object a collection for Q values, the Q value was essentially the thickness of the rewarding lines. Let's say we had a sensor, we had 2 Q values, one to the sensor above it and one to the sensor to the successful path, so that given sensor was set up as a linked hash map where the key was the different sensors around each sensor and its neighbors and the value is the actual Q value that's the thickness and that was updated while the model is training. While the model was on training those line thicknesses also changed. The sensor agent had three functions it could return, the reward for going to a specific other sensor or to a destination sensor. There were two special cases, either the target sensor we wanted to go or to any other sensor that can appear on the way to the destination. We basically hard coded a little penalty is minus 1 times the distance to that destination / 10 so it's typically not 0.1 or something. A small negative value unless the sensor was one of the neighboring sensor to the destination and then moving to that destination gave a huge reward 100 again. There was also a function to retrieve an array list of all feasible neighbors sensors so that was a essentially looping across all neighbor sensors, eight neighbors around any given sensor actually and then the feasible neighbors shift through those eight and the last function was to use that path containing the sensors with highest rewards. It looped through all the all the feasible neighbors in the Q values collection and basically got the one that has the highest Q value.

The Pathfinder class, which was actually going to that sensor it made a decision during the training mode. He updated the q-value of actually going there and the valuable reward, so by actually going somewhere, pathfinder had to reap a reward or it got a little penalty. After leaving, it disconnected itself from the previous sensor it just left and jump to the new sensor and connect to that, update Q value.

The agent classes sensor and pathfinder were implemented on the main and we created a population of sensors in partition, obviously because we needed lots of them and positioned them randomly each sensor found its own location based on some code with the index and we initially assigned all of them to be green sensors then we also did set up one pathfinder so it was a single agent and we just put it on the main because it needed to live. The model did initiate sensor connections a tiny fraction of a second after the model was started then the pathfinder looped across all the sensors, every one of them and then for each sensor we also looped across all its neighbors all the eight neighbor and we also did fill out for every cell in the collection Q values where the key was any neighbor and then initially we just put it at zero and remember that pure value is the reinforcement reward so at the model start going from any sensor to any neighbor sensor is always zero, that means there' would be no line visible as well. Only when we did update Q values to be positive there was a green line path with reward and if we had a negative reward found from one sensor to another, we updated with a little bit line so initially, we set everything all the Q values and connections up and set them to zero and then there was one more event, the update training step event and that was also triggering cyclically every second in our model in real time. We also updated the Q values using the Q value function so it learned about how good was it to go from the current position of the pathfinder to the next one, given the subsequent positions that come after the next one so that was the one thing that this update training step event did and then there was a simple check as if we actually did reach our target, which typically pathfinder did not in most of the cases but in some case actually the sensor that we did move to was our target sensor (near the base station in each field) and in that case we went back to our initial sensor, which we've stored locally and we basically restarted the training, we say we'd done one learning past. The number of passes were also counted and the successfully found paths were stored as well for each sensor individually. We also had a bellman discount factor, if we made it smaller it's basically harder for the model to find the destination and it typically needed more training passes. The training was done individually for each field and each field had a destination sensor. As soon as the path-finding process was complete, we use those paths to send data from each sensor from the field to the base station.

4.4.2 Sensor Cycles

In between there are sensor cycles when sensor could read sleep and transmit its data. For sensor cycles, the time to sleep was kept 3 minutes and time when the sensor had to read its data was set for 10 seconds and transmission time was kept 5 seconds.

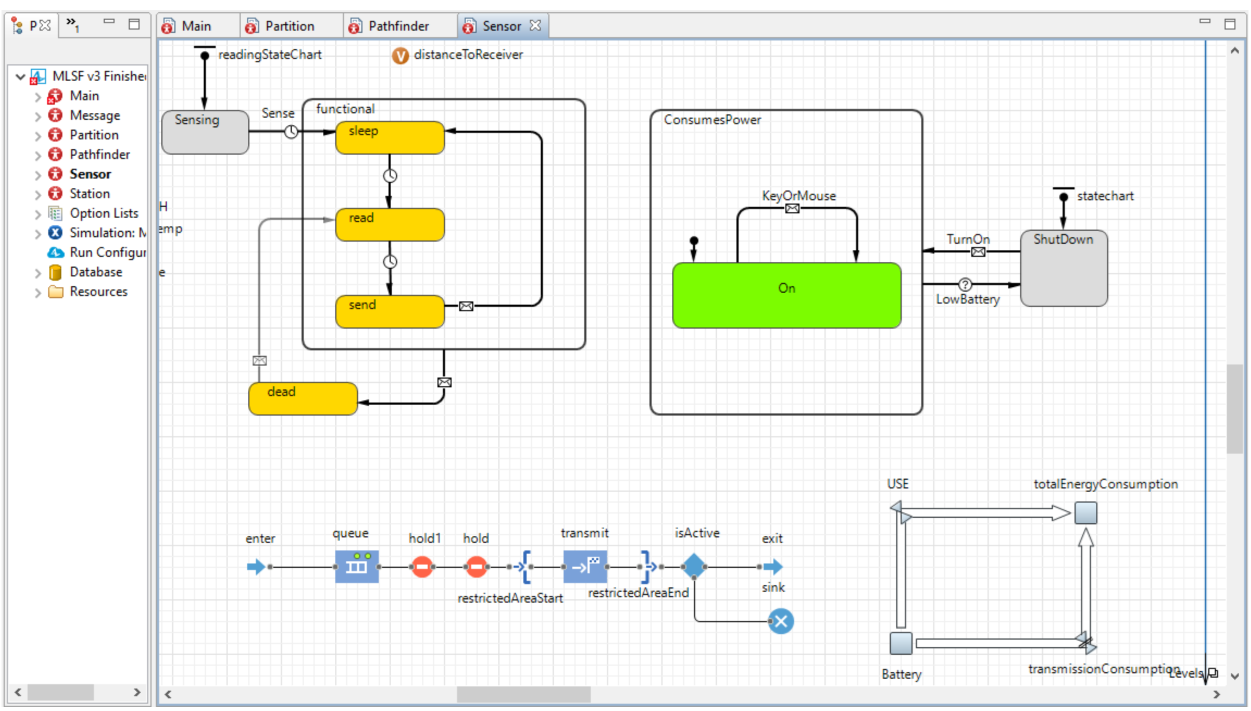


Fig. 7: Sensor cycles state chart

4.4.3 Battery Dynamics

As for as the battery is concerned, we did set battery units for 1000 on generic level which we did change it for different tests. We did use 4 units for reinforcement algorithm power consumption. For sensor different states, 3 units for send state, 2 units for reading state and 0.5 units for sleeping state. The threshold level was set at 10 units and the sensor node was kept dead at less than 4 units left in the battery.

The battery consumption formula was set as:

$$1000 - (\text{USE} + \text{transmission consumption})$$

Where (USE = send +read + sleep) and transmission consumption covers energy consumption during reinforcement learning and transmission of the data.

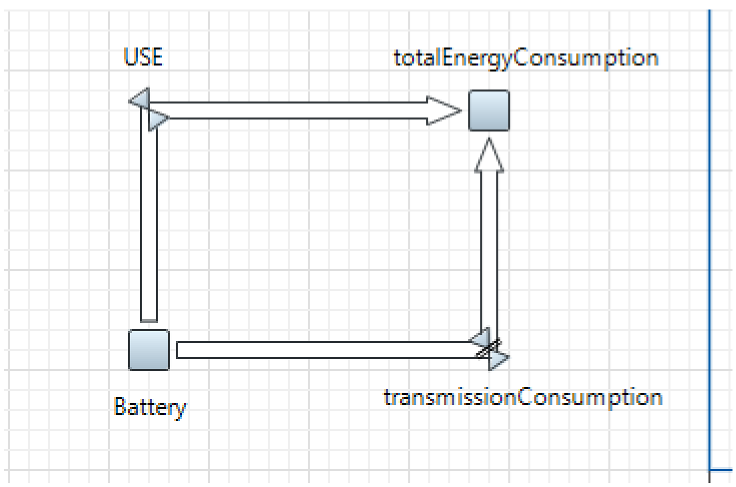


Fig. 8: System Dynamics for battery

Chapter 5: Results and Analysis Discussion

This chapter contains the results from our simulation model and analysis about the information we extracted from the model data. There is also a comparison between our model and two model simulations that we also did simulate in AnyLogic. This chapter contains the topologies and the set of results we did create from them. The later sections of this chapter contain detailed analysis of graphs, images of the topologies and also the comparison data of all three models that we did experiment on.

We actually created three different topologies simulations. First was the broadcast transmission topology, the second one was our topology that used reinforcement learning and the third one was the topology that used shortest path for data transmission. In in all these above mention topologies we did use random deployment of the sensors. More details are given below.

5.1 Parameters used in the model simulation

Here are some common parameters that we did set for all three models to make them equally comparable with each other.

- Battery units were set at 1000 units for all three topologies.
- Model time scale unit was kept in minutes.
- All sensing and reading time units word kept same in all three models that is (sleep time = 3 minutes, read time = 10 seconds and transmission time = 5 seconds.)
- The number of sensors for each field were kept equal in all comparisons of the topologies lying in the same category. Explanation of categories is given below.
- We did take two parameters: average time taken for messages sent and the energy consumption for the messages sent in the lifetime of a network.

5.2 Categories of the Sensor set

In order to compare a lot of sensors in multiple simulations and to observe the behaviour of the field data transmission in different number of sensors we did create some categories where we could observe the transmission behaviour and generate inferences from them. The details of the sensor set is given below.

Number of fields = 6 (in all simulations)

We wanted to do a range of tests for lowest to highest number of sensors for each field and 6 field the number of sensors would be the multiple of 6. For example Set 1: Each field had 22 sensors so 22x 6 = 132 and so on.

	No. of sensors in 1 field	No. of sensors in 6 Fields
Set 1	22	132
Set 2	32	192
Set 3	42	252
Set 4	52	312
Set 5	62	372
Set 6	72	432

Table 5.2.: Sensor Set (Number of sensors in 1 and 6 fields respectively)

5.2.1 Broadcast Transmission

The first model we created for comparison with our proposed solution used broadcasting transmission technique in wireless sensor network. This model followed a distributed networking formation technique where each node had to manage its own information locally. The network formed was self-organized where every node had sent its information to its neighbors to find the target node. The route discovery was done based on the nearest hop available. A mesh network was formed where all sensor nodes were connected with each other finally connecting the base station. Each sensing node forwarded the message to its immediate hop and the message reached the base station.

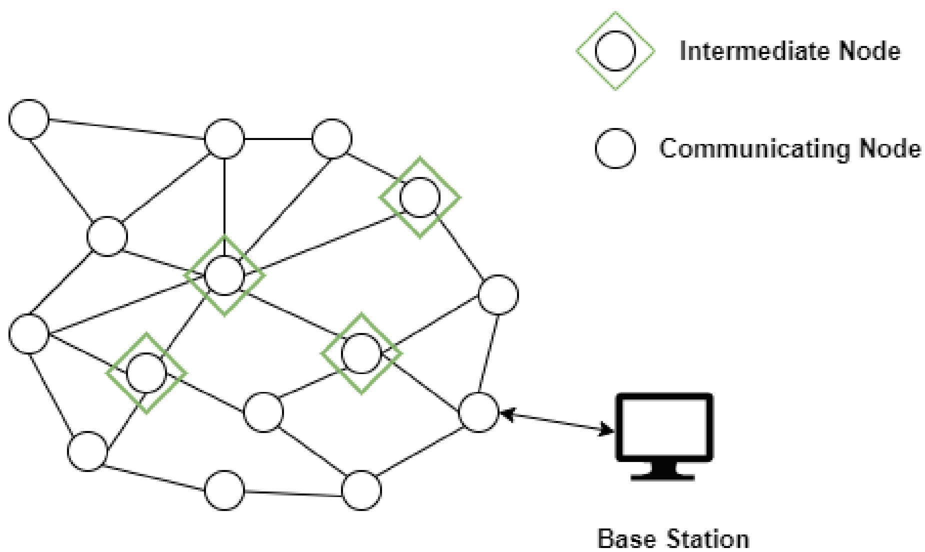


Fig. 9: Block diagram for self-organizing strategy in WSN.

The connected nodes in the broadcasting technique had to keep active most of the time in order to store and forward the message from previous nodes. To optimize the model, the duty cycles of the sensor nodes were set to transmit almost at the same rate in order to keep sensors in sleep state at almost equal time but difference occurred after time (due to difference in distance from neighbors and energy consumption for sensing and processing), in every test in sensor cycles and only successful messages reached at the base station. There was a flood of messages from the sensor nodes to the base station and the number of successful messages was quite good during multiple simulation runs. The base station was meant to have infinite power to record all incoming messages. The broadcast technique model was quite fast to organize itself in the fields as there was not much configuration needed but reorganizing the network was needed when any node dies and the link breakage occurs between the nodes. The network configuration step occurs again for the remaining nodes. This process of self-organizing and transmitting messages over the network repeats itself until all sensor node fall out of energy.

In AnyLogic, we used find nearest node function, geographic coordinates and the farm field coordinates to calculate the distance of nearest node. The nodes then save the information and forward it to their next hop, the process repeats till the information reaches the base station as a target node. After establishing the network, duty cycles of the sensors started working and after each cycle of sleep, read and transmit one successful message reaches the base station, where the information is saved as an excel file with the difference in timestamps of received messages.

After developing the model for broadcast transmission we did several tests on the deployed sensors. Starting from 22 sensors in each field all the way up to 72 sensors in each field for a large scale farming model where the number of fields were kept 6. Average transmission delay and average energy consumption during that transmission for each field is explained with a set of graphs in the following paragraphs.

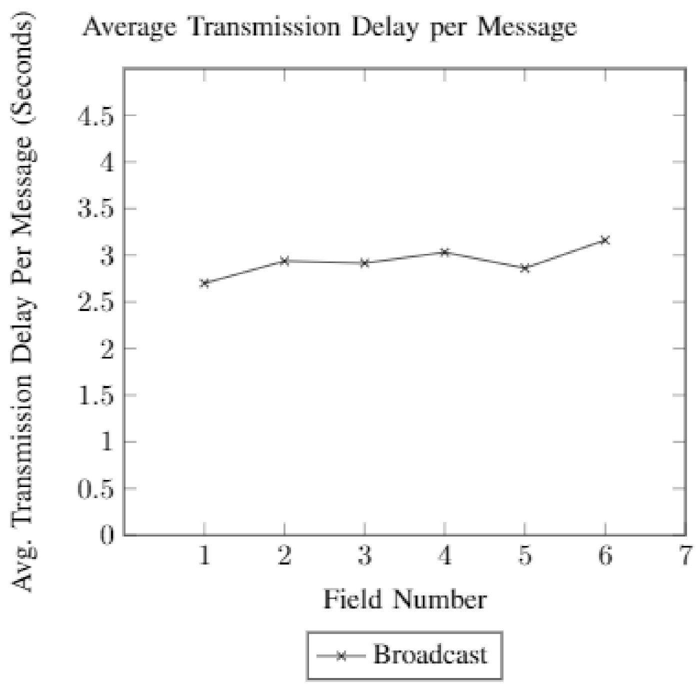
The following screenshot shows the working broadcast transmission technique in AnyLogic, where 6 fields were deployed.

Fig. 9 (a): A random screenshot taken during the executing simulation for Broadcast transmission model

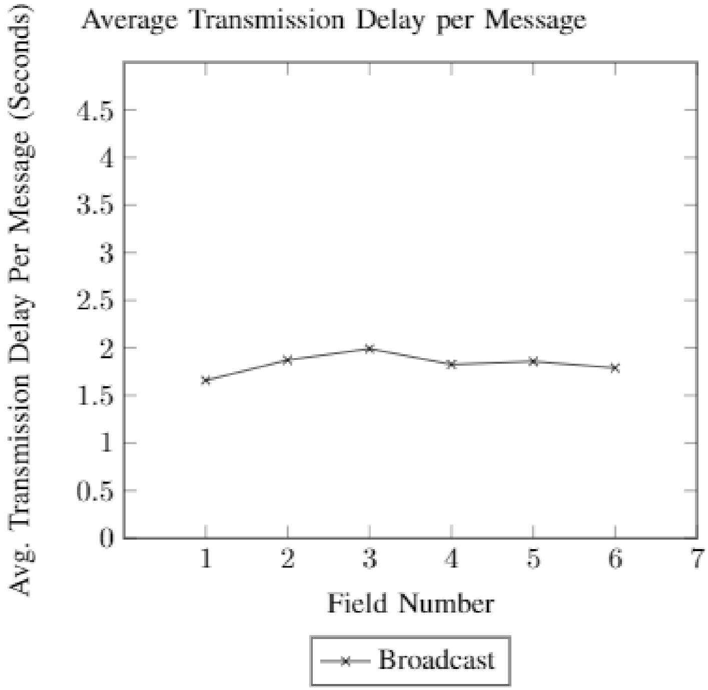
A) Average Transmission Delay per Message in Broadcast Transmission Model

Here in the following set of plotted graphs for broadcast transmission model, on the x-axis there is the number of fields and over the y-axis there is the transmission delay in seconds taken per message for each field. There were nearly 5000 messages and each message was sent after one duty cycle of a sensor so there were almost 5000 number of cycles as well. The average transmission delay was calculated from the received information from the sensors which contained the average time taken to complete a cycle, number of cycles, number of messages.

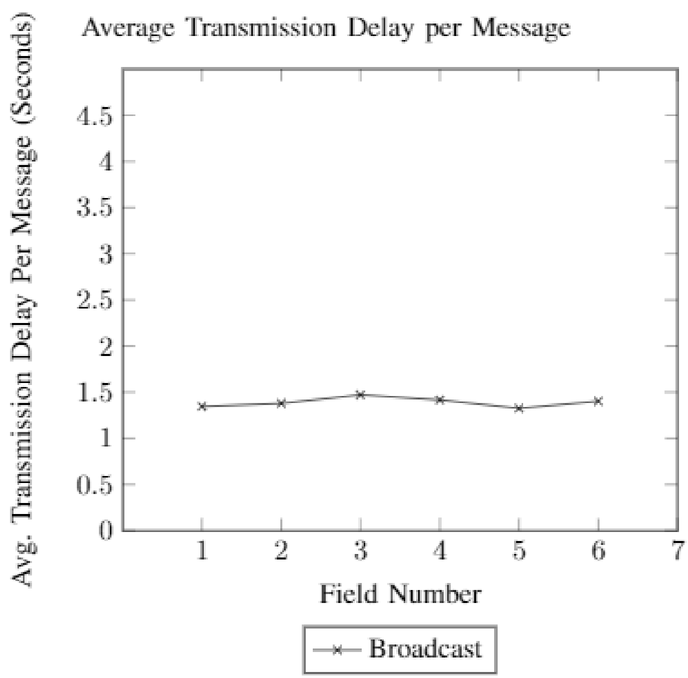
1. The following plotted graph shows the trend for average transmission delay taken for data transmission in large scale where 22 sensors were deployed in each field.



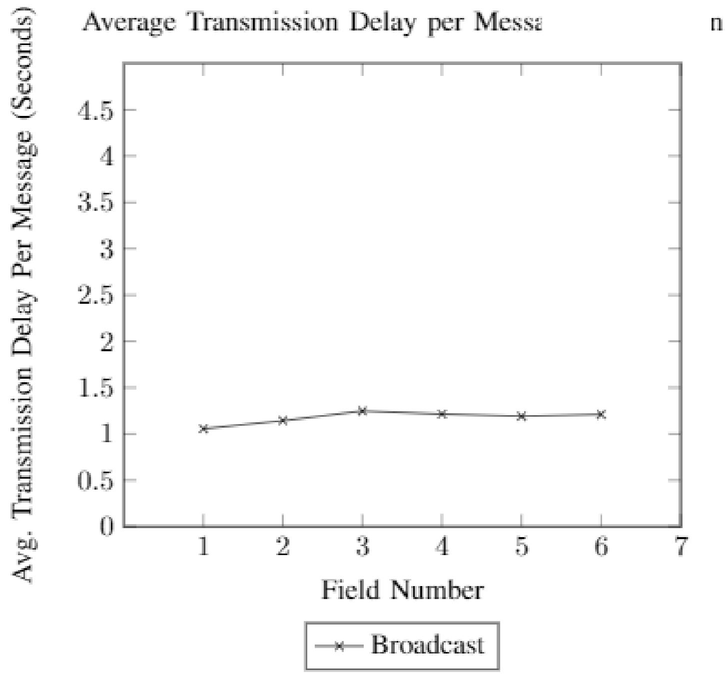
2. The following plotted graph shows the trend for average transmission delay per message transmission in large scale where 32 sensors were deployed in each field.



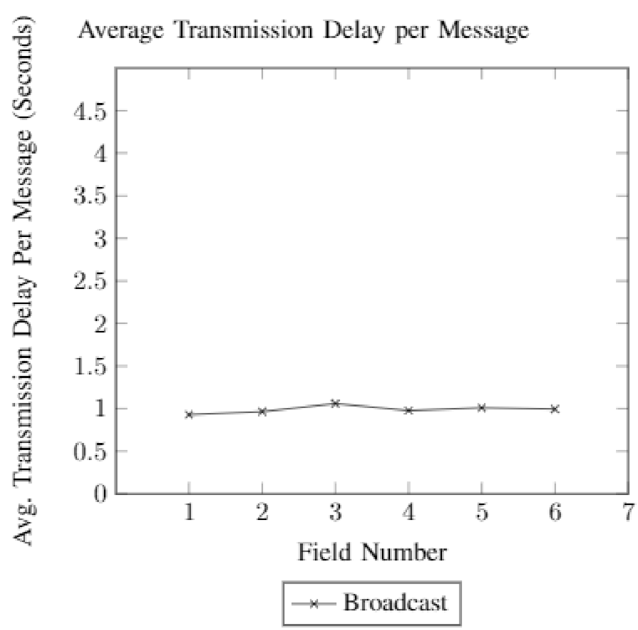
3. The following plotted graph shows the trend per message average transmission delay in large scale where 42 sensors were deployed in each field.



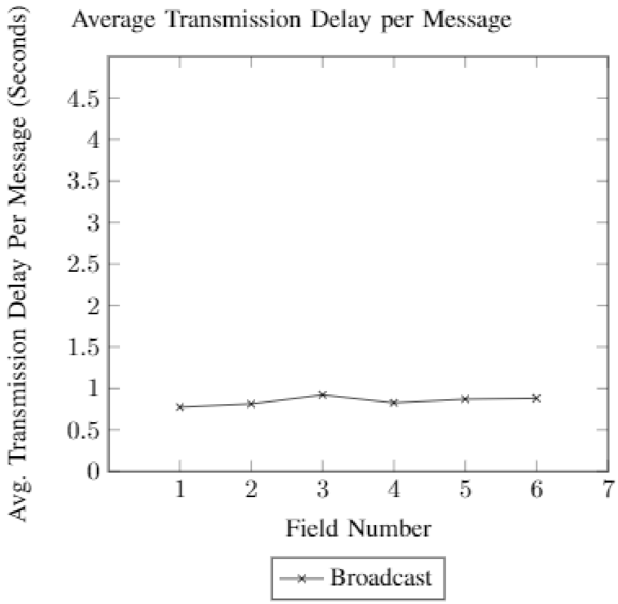
4. The following plotted graph shows the trend for average transmission delay per message taken large scale where 52 sensors were deployed in each field.



5. The following plotted graph shows the trend for average transmission delay taken for one message in large scale where 62 sensors were deployed in each field.



6. The following plotted graph shows the trend for average transmission delay taken for per message data transmission in large scale where 72 sensors were deployed in each field.

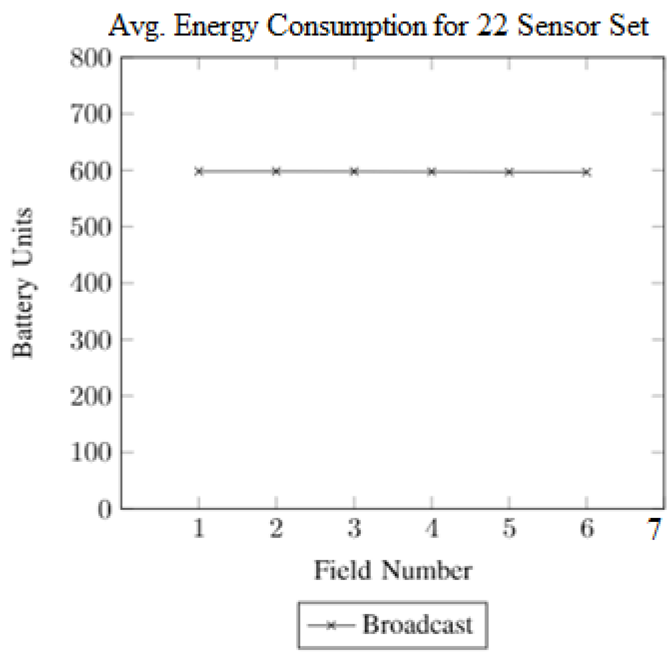


B) Energy Consumption in Broadcast Transmission Model

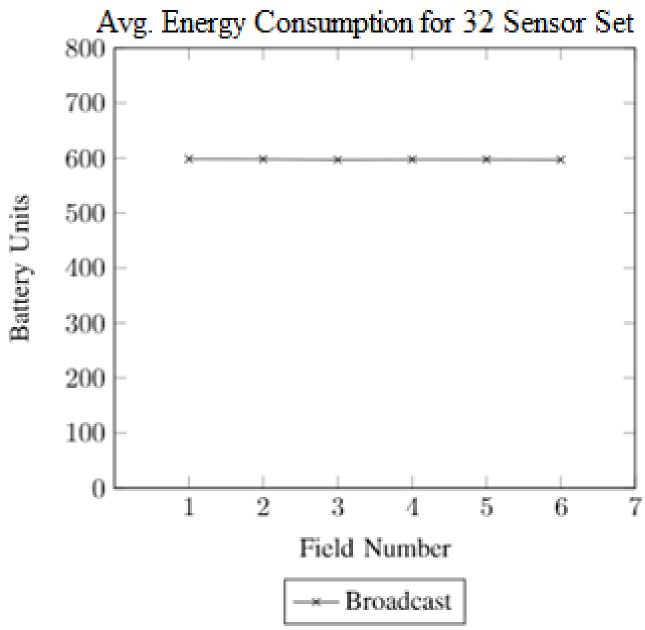
We also did an average energy consumption comparison for each set of sensors in large scale. The plots show the average energy consumption for 5000 messages. The energy consumption data included the energy consumed during the duty cycles. The energy consumption data was taken till the death of the complete network for each set in large scale.

Here in the following set of plotted graphs for broadcast transmission model, on the x-axis there is the number of fields and over the y axis there are battery units consumed by each field.

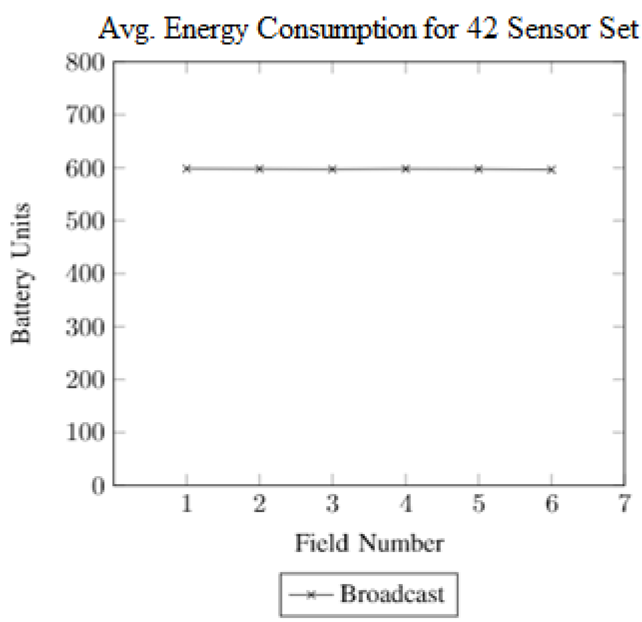
1. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 22 sensors were deployed in each field.



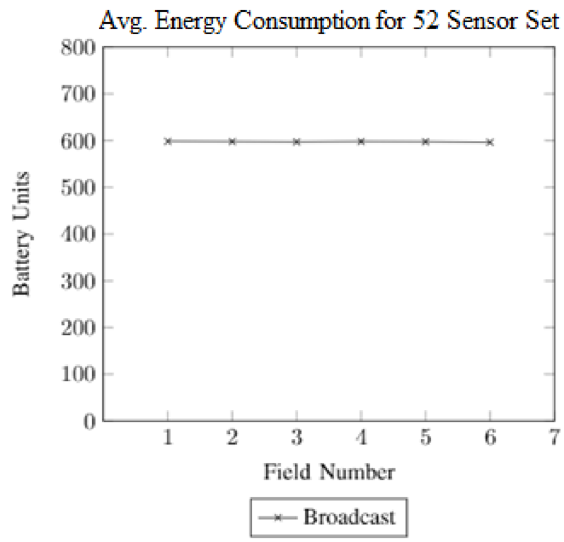
2. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 32 sensors were deployed in each field.



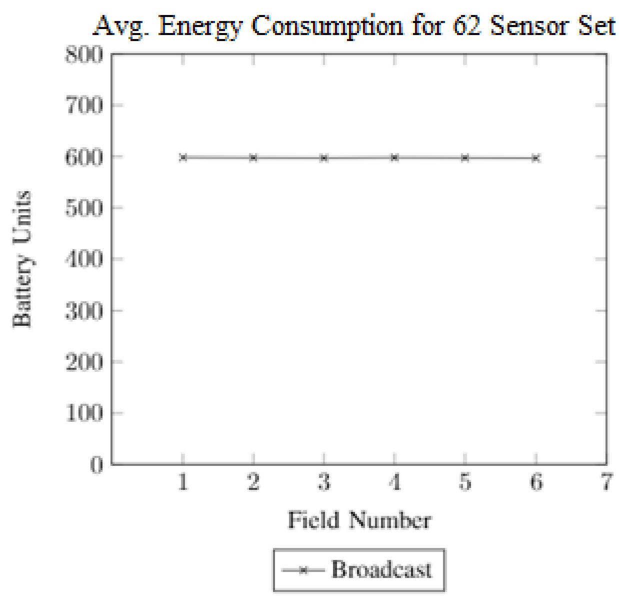
3. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 42 sensors were deployed in each field.



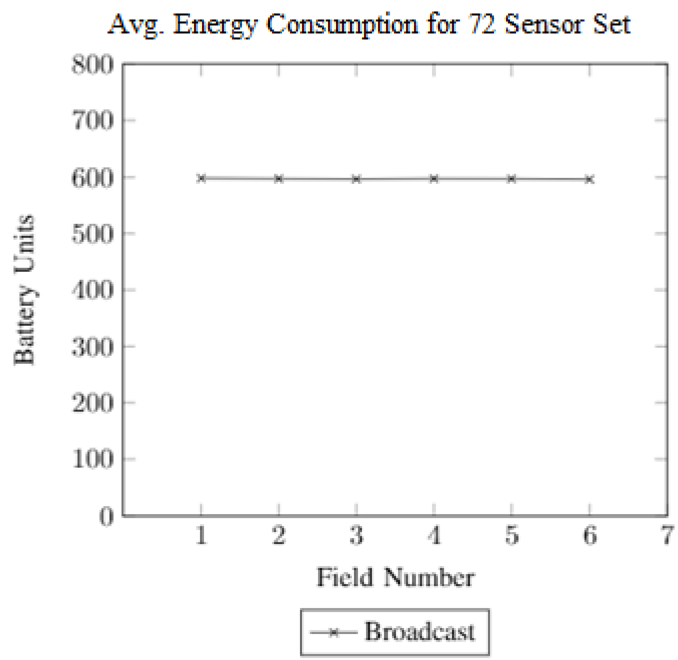
4. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 52 sensors were deployed in each field.



5. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 62 sensors were deployed in each field.



6. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 72 sensors were deployed in each field.



5.2.2 Shortest Path Transmission

Another model we did create for comparison was the shortest path transmission. This model used the cluster head method for network organization. A cluster head organization was chosen among the nodes which was then used to transmit the messages and reorganize the network if any node died.

In Clustering model, sensor nodes are deployed randomly in each other's range, a cluster head is chosen based on some parameter, either remaining energy or the distance from the sink etc., The cluster head is responsible for referencing the sensor nodes to transmit their data in order to forward it to the sink node over the number of hops and in case of any network change happening, the cluster head reorganizes the connections with other nodes.

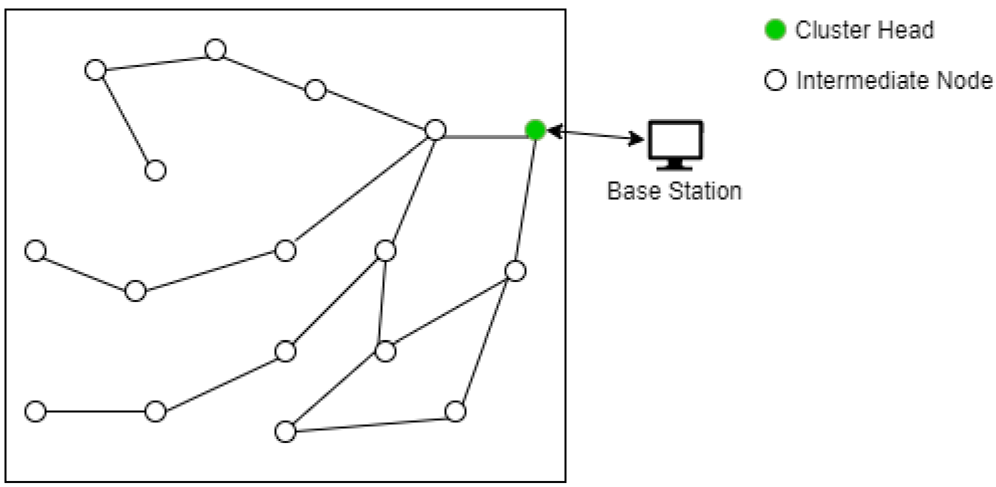


Fig. 10: Block diagram for clustering used in our work.

The figure above represents the clustering model that we used to organize our network to find the shortest paths. We used one farm field as a one cluster where we created a Cluster head from the nodes deployed based on the shortest distance from the base station. AnyLogic has a function to limit the agent connections within a specific area or another agent (farm in our case). Based on the node coordinates, we used find shortest distance of any node from the base station agent which itself selects the node having shortest distance from the base station and considers it as a cluster head. Other nodes use it as a reference for finding the shortest distance for connection through multihop connections. After the successful path organization has been done, the node start using those paths for data transmission.

For each field there is a separate cluster and clustering process occurs to find the cluster head having the shortest distance from the far away base station based on its geographic coordinates. The cluster heads forward the data from its cluster to a nearest node in another cluster where the data is forwarded to its cluster head through the established paths gradually towards the base station. When a node dies clustering process repeats itself finding the shortest paths towards the base station accommodating the cluster head and the intermediate nodes.

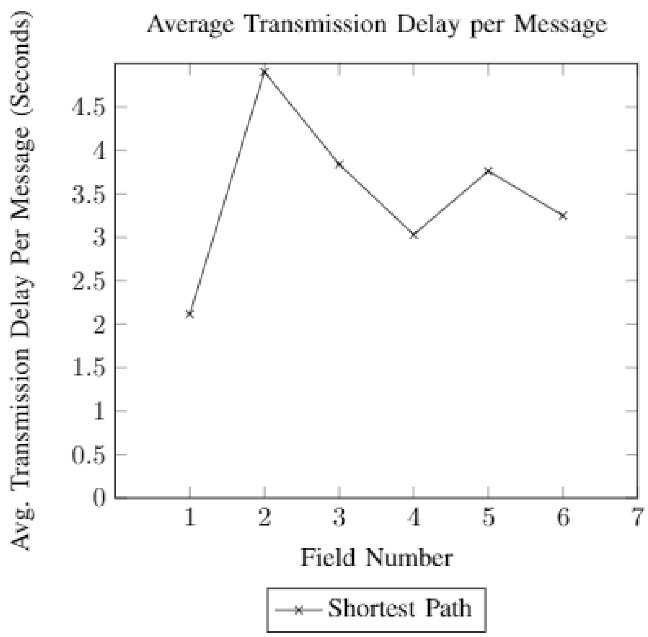
We used shortest path model for testing a number of sensors. The sensor categories were kept same for all models in our research so shortest path transmission also got the same set of sensors in its simulation tests. The sensing parameters were kept same for this model as well and we made an average delay during message transmission and we also made an energy consumption for the sensor set used for shortest path transmission. The following paragraphs explain the results and discussion about observations we gathered from the model data.

Fig. 10 (a): A random screenshot taken during the executing simulation for Shortest Path transmission model

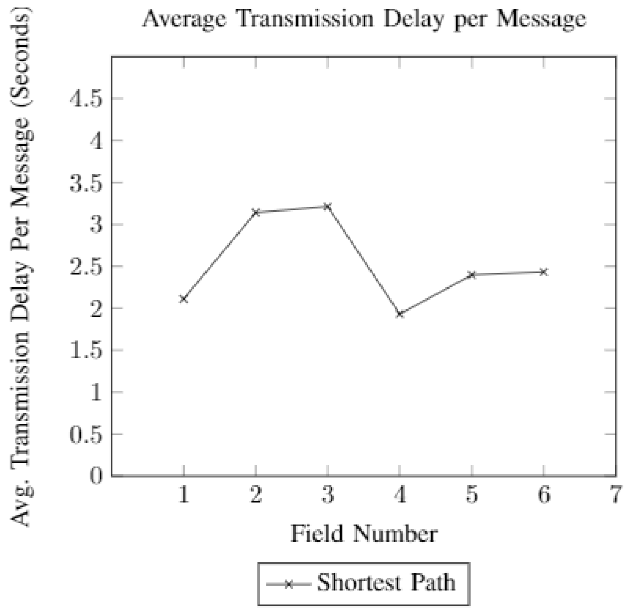
A) Average Transmission Delay Taken in Shortest Path Transmission

Here in the following set of plotted graphs for shortest path transmission model, on the x-axis there is the field number and over the y axis there is the time in seconds taken for each field.

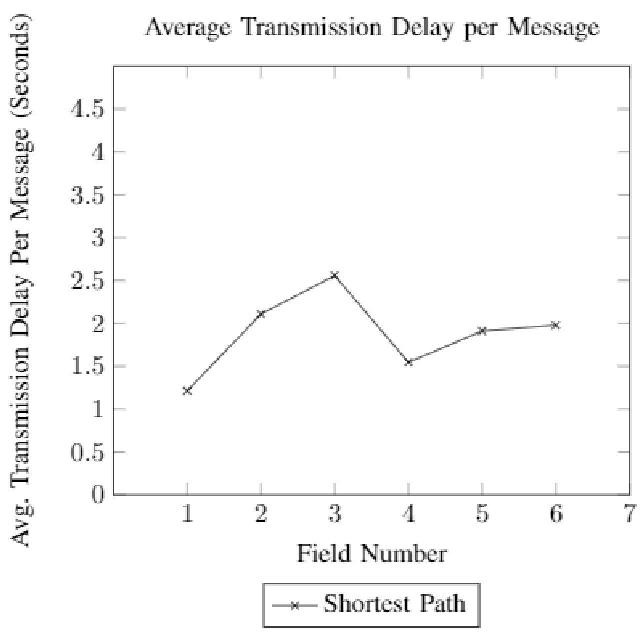
1. The following plotted graph shows the trend for average time taken for transmission delay in large scale where 22 sensors were deployed in each field.



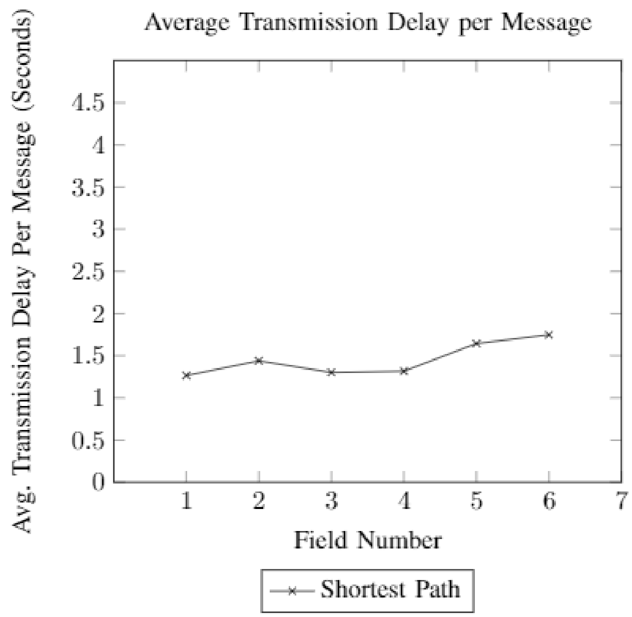
2. The following plotted graph shows the trend for average time taken for data transmission in large scale where 32 sensors were deployed in each field.



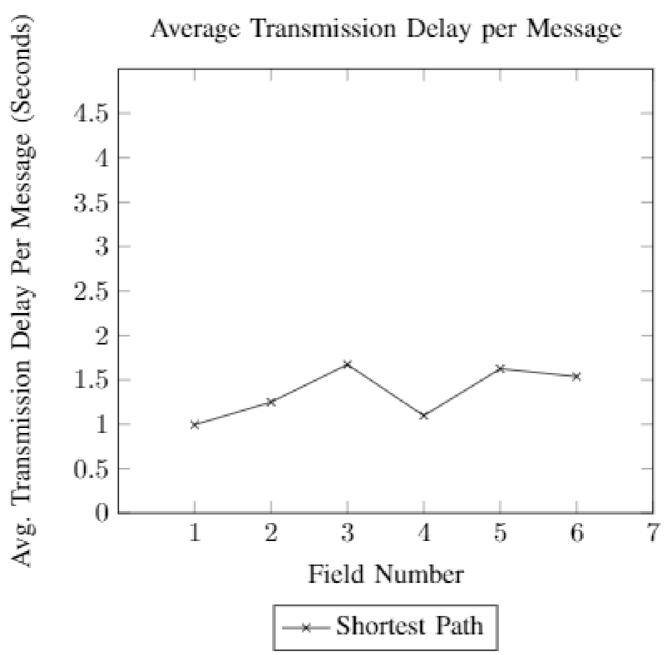
3. The following plotted graph shows the trend for average time taken for data transmission delay in large scale where 42 sensors were deployed in each field.



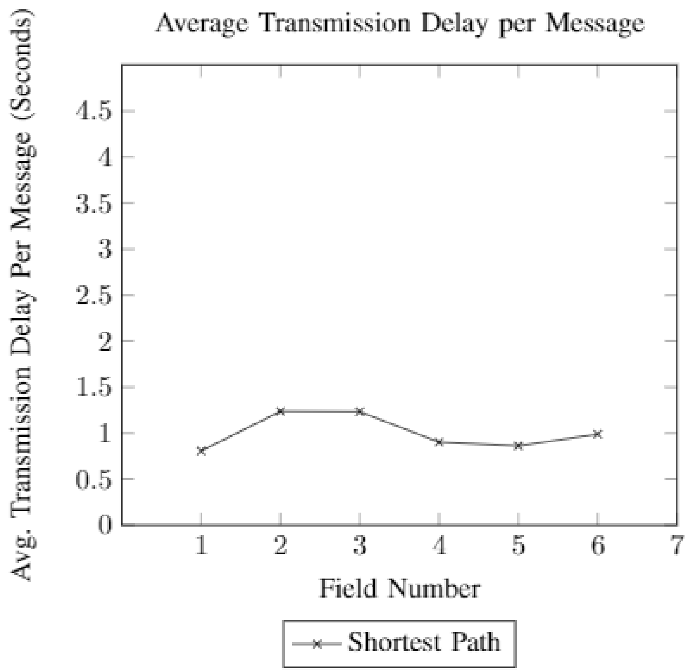
4. The following plotted graph shows the trend for average time taken for data transmission delay in large scale where 52 sensors were deployed in each field.



5. The following plotted graph shows the trend for average time taken for data transmission in large scale where 62 sensors were deployed in each field.



6. The following plotted graph shows the trend for average time taken for data transmission in large scale where 72 sensors were deployed in each field.

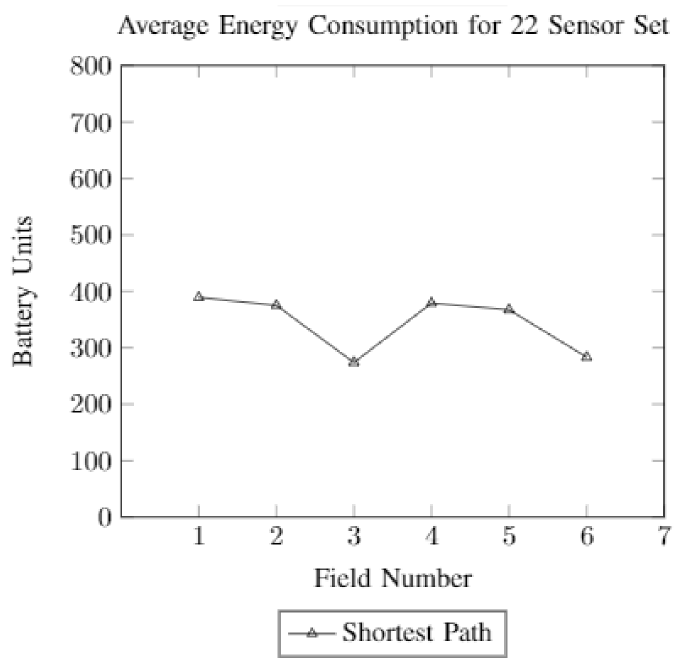


B) Average Energy Consumption Comparison for Shortest Path Transmission

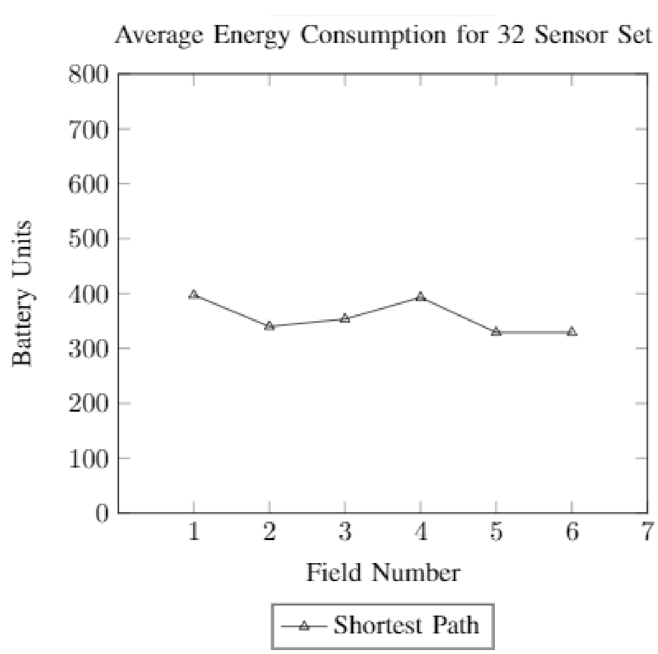
We also did an average energy consumption comparison for each set of sensors in large scale. The energy consumption data was taken till the death of the complete network for each set in large scale.

Here in the following set of plotted graphs for shortest path transmission model, on the x-axis there is the field number and over the y axis there are battery units consumed by each field.

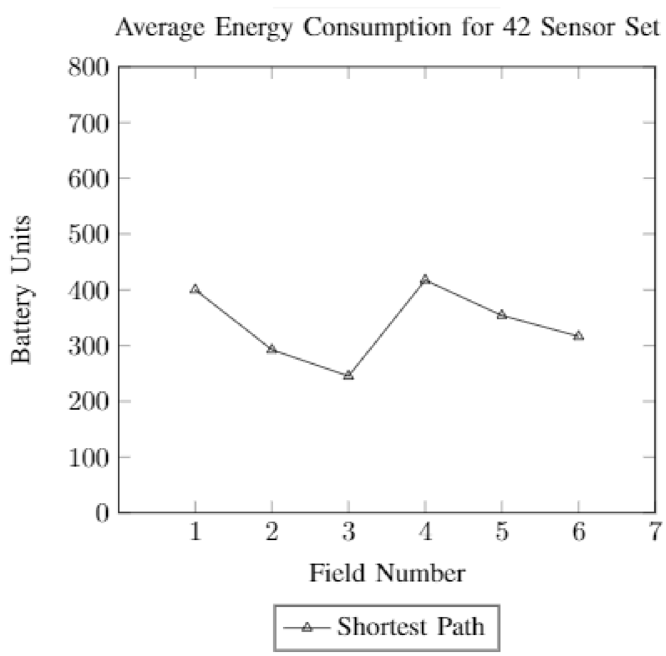
1. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 22 sensors were deployed in each field.



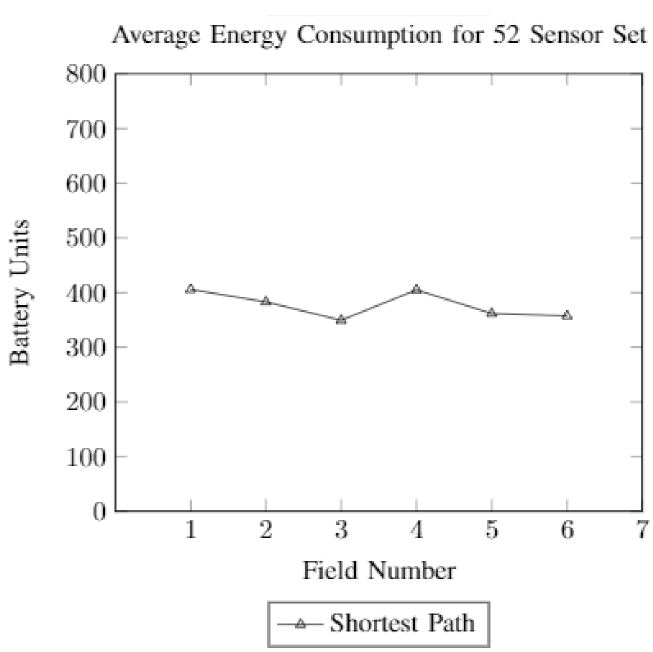
2. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 32 sensors were deployed in each field.



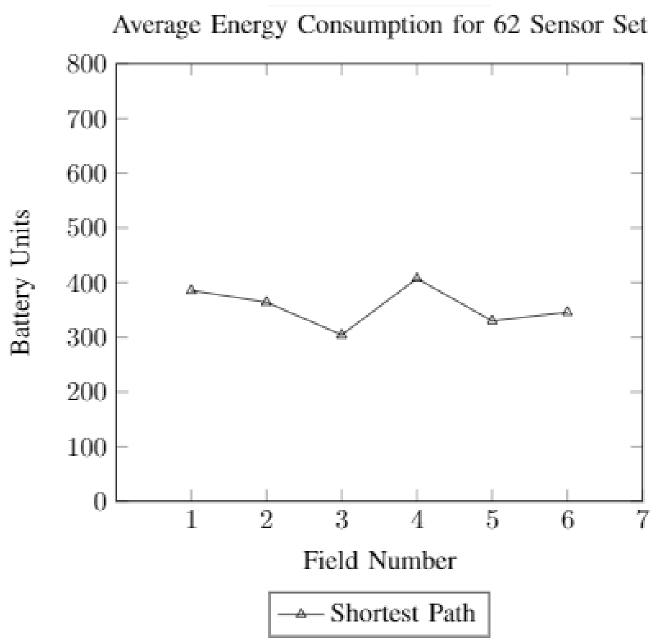
3. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 42 sensors were deployed in each field.



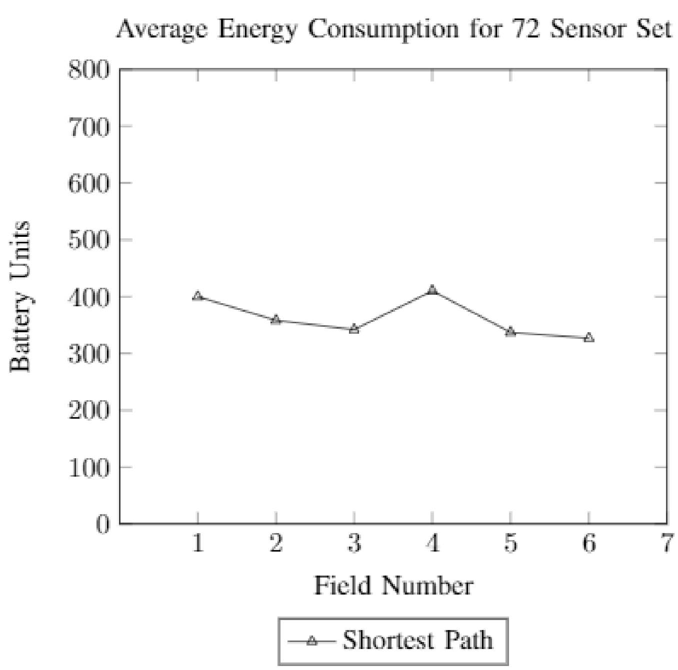
4. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 52 sensors were deployed in each field.



5. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 62 sensors were deployed in each field.



6. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 72 sensors were deployed in each field.



5.2.3 Transmission Using Reinforcement Learning

The third and the final model we did make was a simulation using the paths created by reinforcement learning. The working of the complete model has been described in the section 4.4.1 and its following sections.

To summarize the working of the model, optimal set of paths towards the base station were found based on the geographic coordinates of sensor nodes. Each field has an edge node which has the shortest distance from the base station which was found using the find nearest function in AnyLogic. The edge serves as the cluster head of a field or the nodes deployed in the field. The edge node runs a reinforcement learning algorithm to find a set of optimal paths and successful paths' state are shared with other nodes to transmit messages over the paths. The edge nodes forward the data to other nodes in other fields and through multihop transmission gradually moving the data towards the base station which lies at the end of complete farm line.

When a node dies its state is removed from the network and alternate available paths are used to transmit messages until all nodes die the network. Several tests have been conducted on different number of deployed sensors grouped into categories. The categories were kept same for all models in our research so the transmission using the paths developed by reinforcement learning also got the same set of sensors in its simulation tests. The sensing parameters were kept same for this model as well and we made an average time taken for message transmission and we also made an energy consumption for the sensor set used for transmission using paths developed by reinforcement learning. The following paragraphs explain the results and discussion about observations we gathered from the model data.
Image has been shown on the next page.

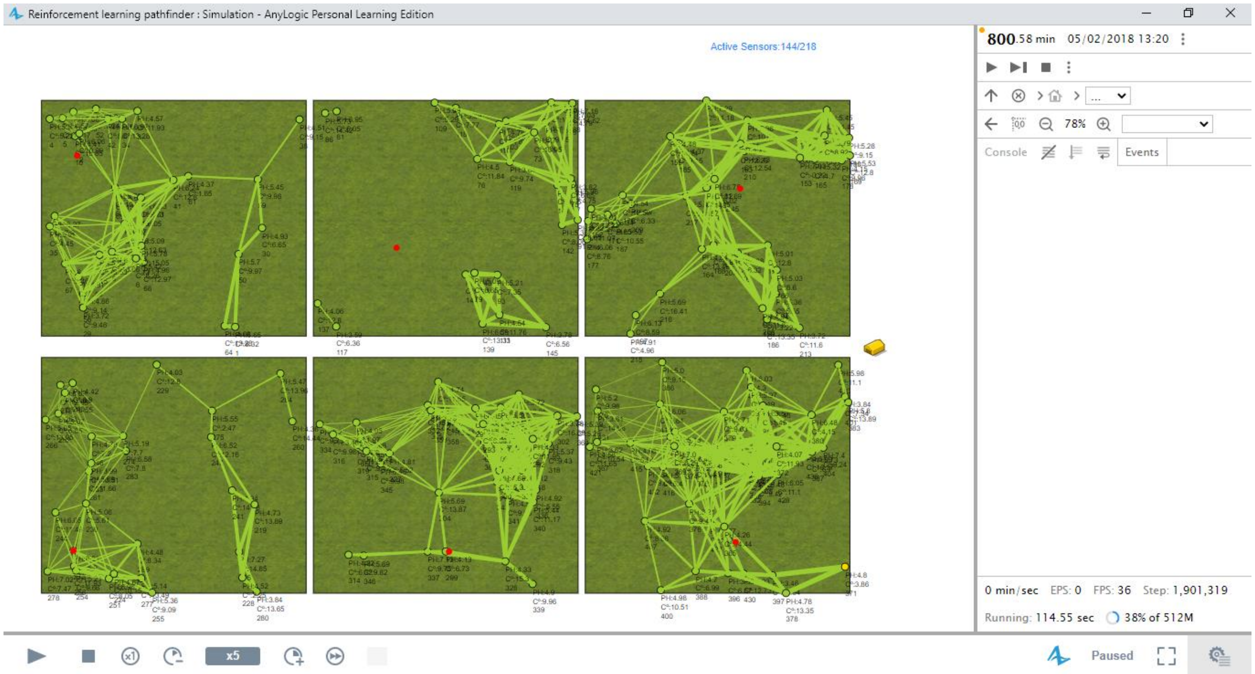
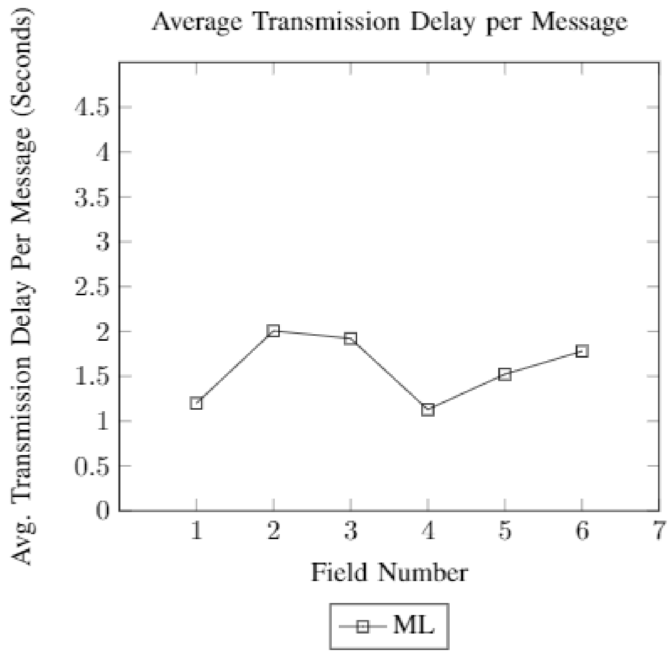


Fig. 11: A random screenshot taken during the executing simulation for the transmission using paths developed by reinforcement learning

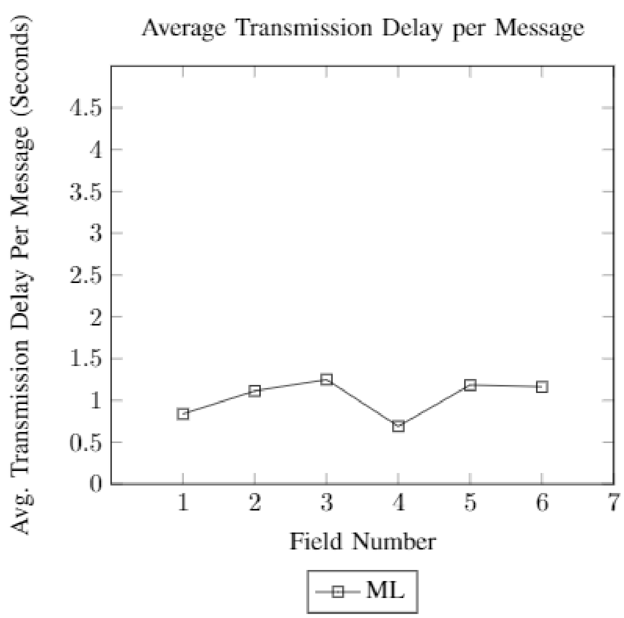
A) Average Time for the Transmission Delay Using the Paths Developed by Reinforcement Learning

Here in the following set of plotted graphs for the transmission delay using the paths developed by reinforcement learning, on the x-axis there is the number of fields and over the y axis there is the transmission delay in seconds taken for each field.

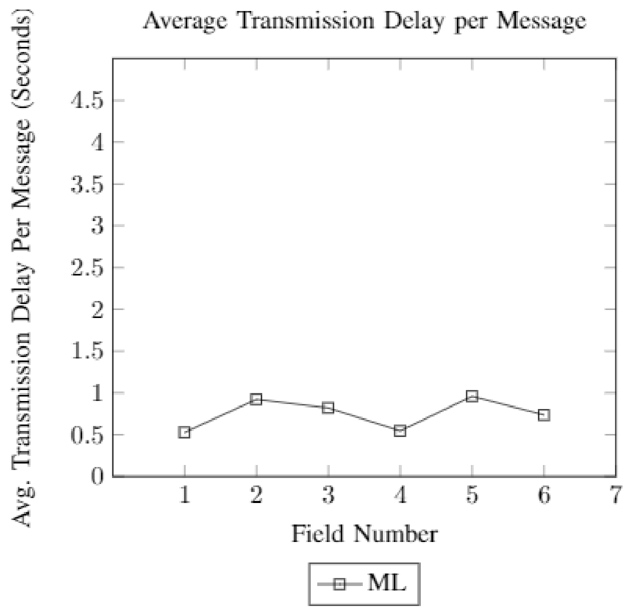
1. The following plotted graph shows the trend for average transmission delay taken for data transmission in large scale where 22 sensors were deployed in each field.



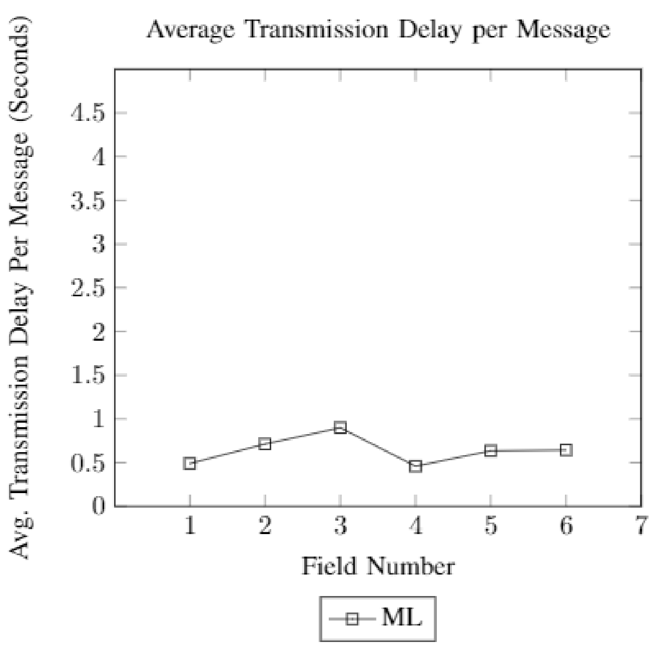
2. The following plotted graph shows the trend for average transmission delay taken for data transmission in large scale where 32 sensors were deployed in each field.



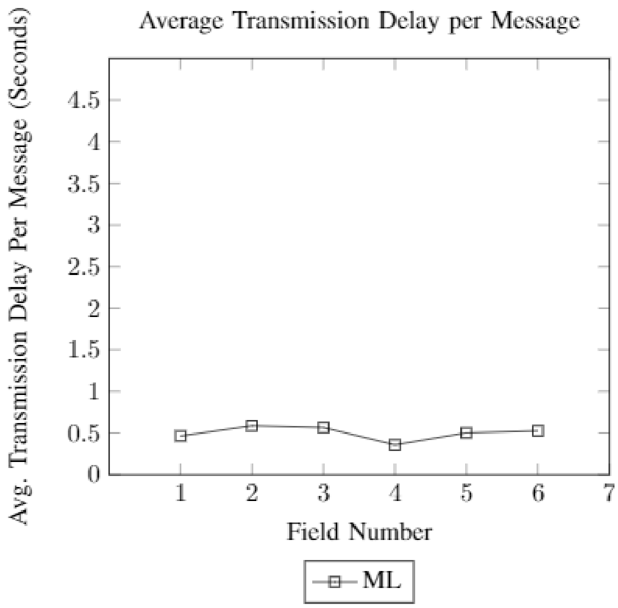
3. The following plotted graph shows the trend for average time taken for data transmission delay in large scale where 42 sensors were deployed in each field.



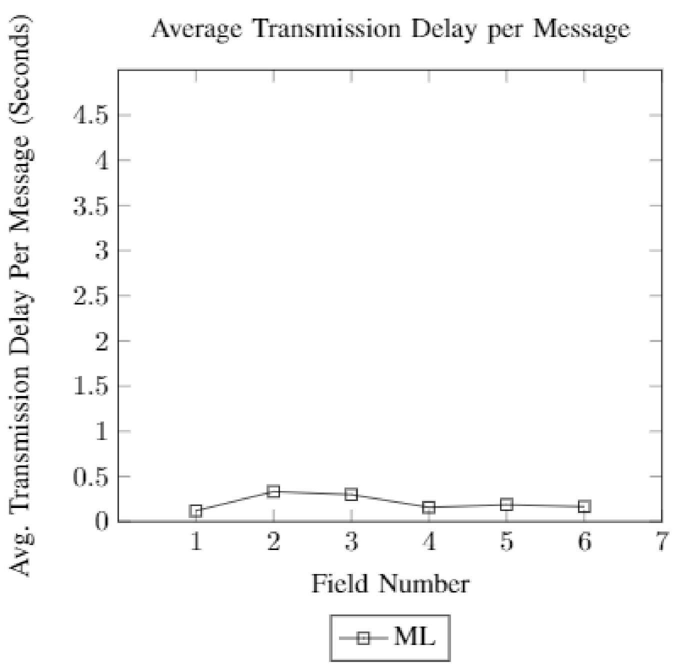
4. The following plotted graph shows the trend for average time taken for data transmission delay in large scale where 52 sensors were deployed in each field.



5. The following plotted graph shows the trend for average time taken for data transmission in large scale where 62 sensors were deployed in each field.



6. The following plotted graph shows the trend for average time taken for data transmission in large scale where 72 sensors were deployed in each field.

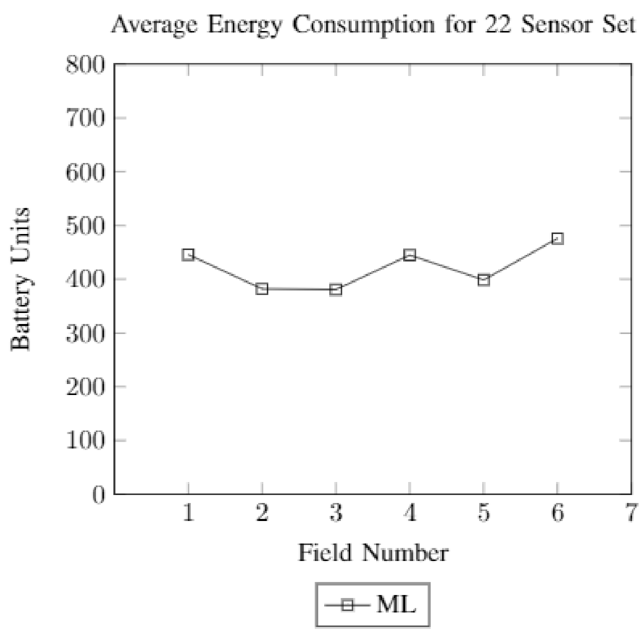


C) Average Energy Usage Comparison in Transmission Using Paths Created from Reinforcement Learning

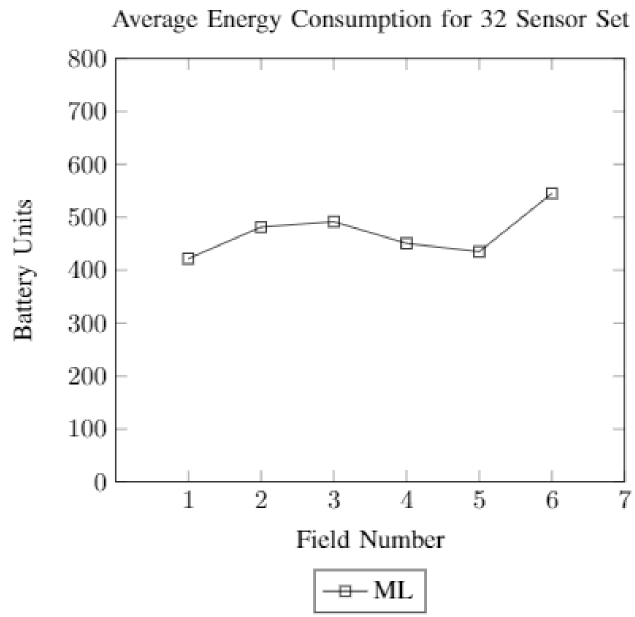
We also did an average energy consumption comparison for each set of sensors in large scale. The energy consumption data was taken till the death of the complete network for each set in large scale.

Here in the following set of plotted graphs for the transmission done using paths created through reinforcement learning, on the x-axis there is the number of fields and over the y axis there are battery units consumed by each field.

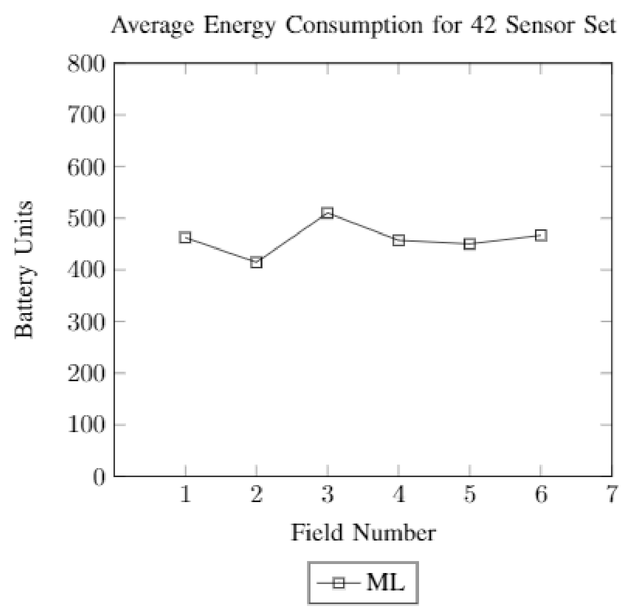
1. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 22 sensors were deployed in each field.



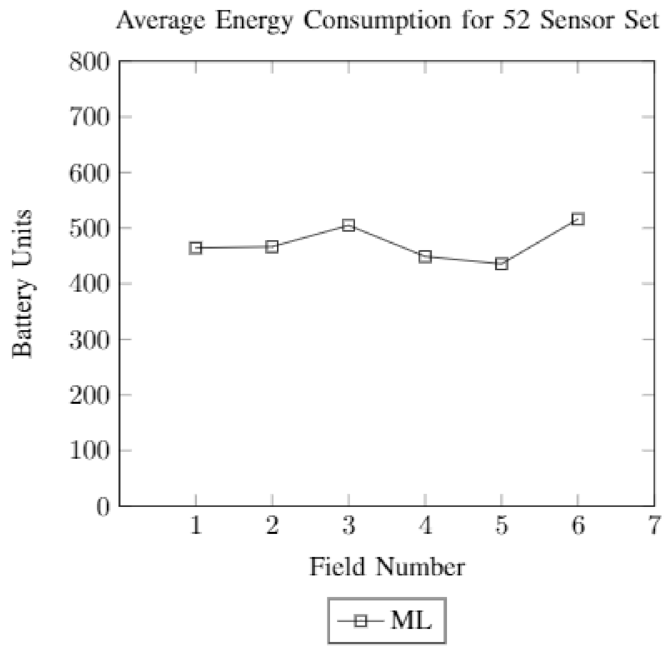
2. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 32 sensors were deployed in each field.



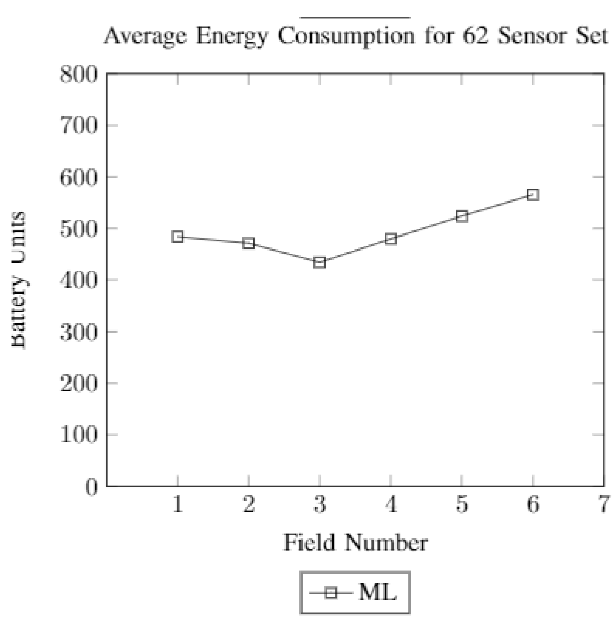
3. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 42 sensors were deployed in each field.



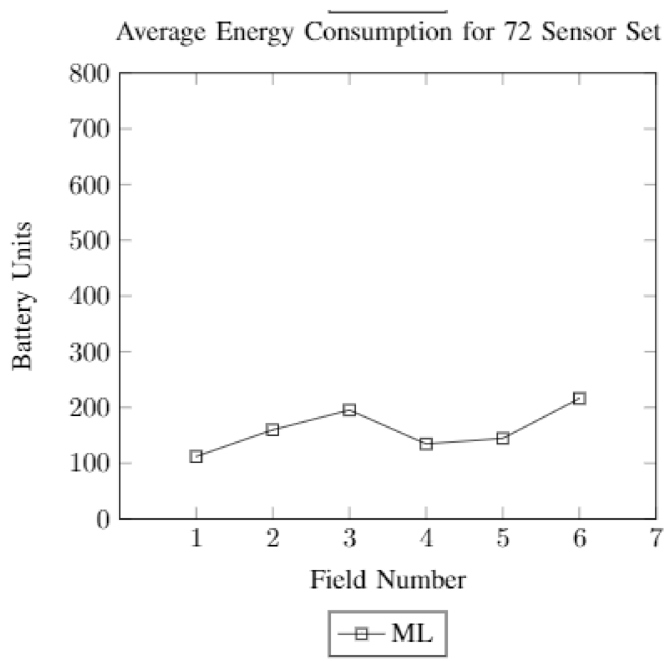
4. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 52 sensors were deployed in each field.



5. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 62 sensors were deployed in each field.

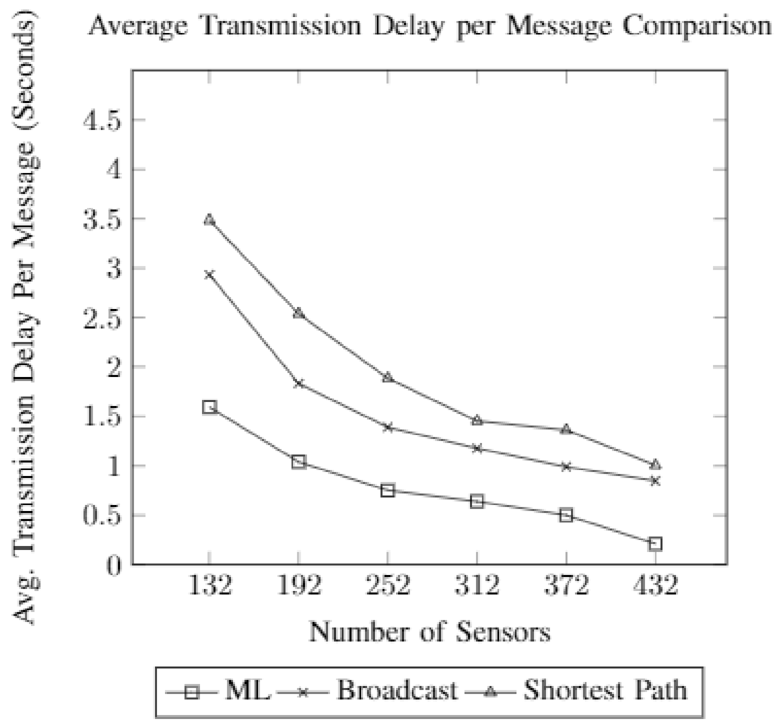


6. The following plotted graph shows the trend for average energy consumption for data transmission in large scale where 72 sensors were deployed in each field.



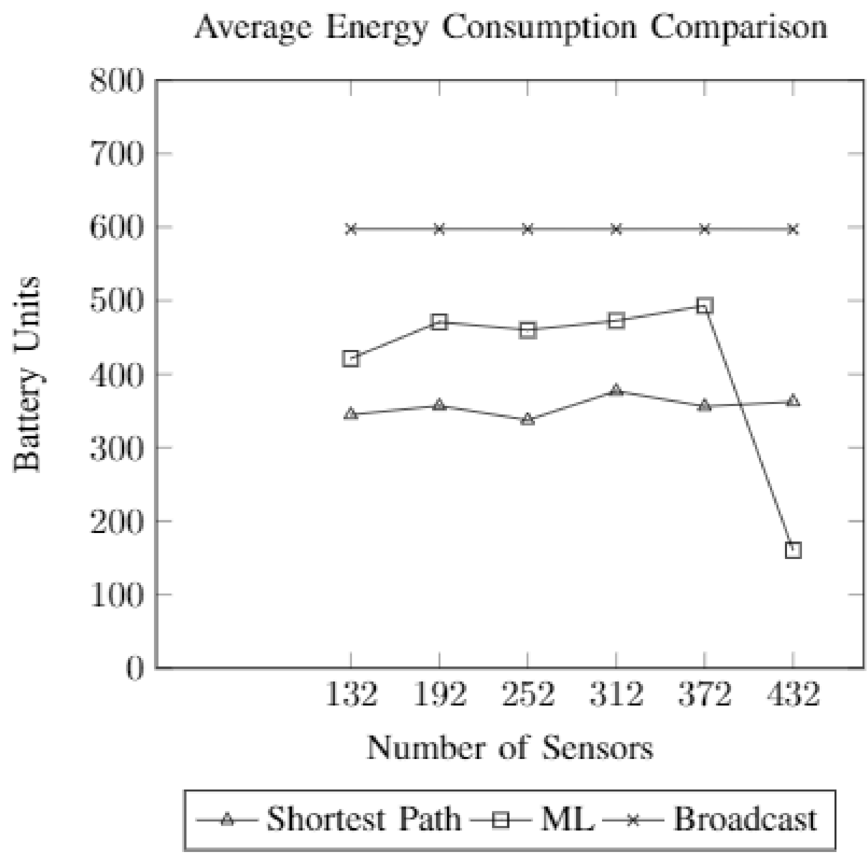
5.3 Average Time Consumption Comparison for all Transmission Techniques

Below is the comparison for the average time taken by different sensor sets for sending messages. On the x-axis there is number of deployed sensors and over the y-axis there is average transmission delay in minutes taken for each message. It can easily be seen that average transmission time per message in reinforcement model is the least and our model is quickest in transmitting for large number of nodes or on larger scale farm networks.



5.4 Average Energy Consumption Comparison for all Transmission Techniques

In the graph given below, on the x-axis there is the number of deployed sensors and over the y-axis there are battery units consumed by each set of sensors. It can be seen from the comparison graph that the increasing the number of sensors decreases the overall energy consumption of the sensor network. The reason is that the large number of nodes had a large set of paths developed for transmission.



Chapter 6: Thesis Conclusion and Future Work

6.1 Conclusion

Our main goal to be pursued in this research work was to develop an efficient transmission system in terms of energy and time. Efficient energy consumption is an emerging area of research. The proposed system is the energy efficient transmission of sensory data in large scale farming by using paths created through reinforcement learning.

The objectives of the proposed transmission are energy consumption minimization, reduction in average time for transmission, improve path finding process during wireless sensor network establishment, to reduce common wireless sensor network problem like energy hole problem and hidden node problem as well as to optimize the network load by using alternate paths available, and ultimately reduce the management process for large scale wireless sensor network especially by increasing the network lifetime. The proposed model is based on reinforcement learning that uses random deployment topology for sensors as an input parameter and creates several optimal paths for data transmission basically to reduce the energy consumption in a wireless sensor network meant for IOT in agriculture on a large scale.

Chapter (2) has shared a generic background and overview of the past research work done in related to wireless sensor networks in agricultural domain. The wireless sensor networks meant for farming are presented in details accompanied with their implementation details of different models. Some of them were not implemented at all. Networking and their efficiency regarding time and energy were also discussed. In the later sections Machine Learning was introduced with its further types and introduction to machine learning in IOT in agriculture are also discussed in this chapter.

In chapter (3), problem statement for our research work is discussed in detail. Introduction to Reinforcement learning along with its types are discussed in detail. The steps of proposed methodology along with reinforcement learning are also discussed in this chapter. In chapter (4), we discussed the methodology used in our research, implementation details of proposed system along with the tools that were involved in the making of the model are discussed. Proposed model with its design and implementation details was also discussed. In chapter (5), performance of data transmission using paths created through reinforcement learning was compared with broadcasting and shortest path data transmission, and the comparison results show that proposed model is 16 to 25% efficient than existing transmission models in average time usage. For energy efficiency, shortest proved to be better for smaller and medium scaled agricultural wireless sensor networks but for larger scales our model out-performed the other two transmission methods in comparison. The results with plotted graphs created on per field level and with different sensor sets were discussed and compared in detail.

6.2 Future Work

Here is the list of the work that can be considered as the potential in this research and some of the work will also be done in the future:

- Implement this system in large scale farming network accompanied with artificial intelligence would be helpful in increasing efficiency.
- Implementing and experimenting more topologies to test energy efficiency.
- The proposed model also has a room for optimization in terms of how data is being processed including the sensor cycles.
- Make this model adaptive in IOT in agriculture after added work done.
- The model can also be extended with little modifications for ad hoc networks and vehicular networks.

References

[1] Malik AW, Rahman AU, Qayyum T, Ravana SD. Leveraging fog computing for sustainable smart farming using distributed simulation. IEEE Internet of Things Journal. 2020 Jan 17;7(4):3300-9.

[2] Ramson, Jino. (2017). Applications of wireless sensor networks — A survey. 10.1109/ICIEEIMT.2017.8116858.

[3] Zhang, C., Zhang, M., Su, Y., Wang, W.Smart home design based on ZigBee wireless sensor networkProceedings of the 7th International ICST Conference on Communications and Networking in China (CHINACOM '12)August 2012463466.

[4] Khan, A. A., Mouftah, H. T.Energy optimization and energy management of home via web services in smart gridProceedings of the IEEE Electrical Power and Energy Conference (EPEC '12)October 2012Ontario, CanadaIEEE141910.1109/EPEC.2012.6474940.

[5] Lim, Y, Kim, HM, Kang, S. A design of wireless sensor networks for a power quality monitoring system. Sensors 2010; 10(11): 9712–9725.

[6] Li, Li & Xiaoguang, Hu & Ke, Chen & Ketai, He. (2011). The applications of WiFi-based Wireless Sensor Network in Internet of Things and Smart Grid. 789-793. 0.1109/ICIEA.2011.5975693.

[7] Shaikh A, Pathan S. Research on wireless sensor network technology. International Journal of Information and Education Technology. 2012 Oct 1;2(5):476.

[8] Kalaivani T, Allirani A, Priya P. A survey on Zigbee based wireless sensor networks in agriculture. In3rd International Conference on Trendz in Information Sciences & Computing (TISC2011) 2011 Dec 8 (pp. 85-89). IEEE.

[9] LoRa Alliance. White Paper: A Technical Overview of Lora and Lorawan; The LoRa Alliance: San Ramon, CA, USA, 2015.

[10] Engmann F, Katsriku FA, Abdulai JD, Adu-Manu KS, Banaseka FK. Prolonging the lifetime of wireless sensor networks: a review of current techniques. Wireless Communications and Mobile Computing. 2018 Aug 15;2018.

[11] Dietrich I, Dressler F. On the lifetime of wireless sensor networks. ACM Transactions on Sensor Networks (TOSN). 2009 Feb 11;5(1):1-39.

[12] Zhang X, Lu X, Zhang X. Mobile wireless sensor network lifetime maximization by using evolutionary computing methods. Ad Hoc Networks. 2020 Apr 15;101:102094.

[13] Ramson SJ, Moni DJ. Applications of wireless sensor networks—A survey. In2017 international conference on innovations in electrical, electronics, instrumentation and media technology (ICEEIMT) 2017 Feb 3 (pp. 325-329). IEEE.

[14] Y. Wang and L. Wang, Application Time Series Analysis, China Renmin University Press, 2005.

[15] X. Xu, Research and application on time series prediction based on data mining method, 2011.

[16] T. Le, Y. Cai, X. Ma and L. Wang, "Development and application of time series forecasting", Ordnance Industry Automation, no. 2, pp. 63-68, 2015.

[17] L. Li, Y. Wu, Y. Ou, Q. Li, Y. Zhou and D. Chen, "Research on machine learning algorithms and feature extraction for time series," 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Montreal, QC, Canada, 2017, pp. 1-5, doi: 10.1109/PIMRC.2017.8292668.

[18] https://www.dataversity.net/a-brief-history-of-machine-learning/#

[19] Goldberg DE, Holland JH. Genetic algorithms and machine learning.

[20] Cunningham P., Cord M., Delany S.J. (2008) Supervised Learning. In: Cord M., Cunningham P. (eds) Machine Learning Techniques for Multimedia. Cognitive Technologies. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-75171-7_2

[21] Hastie T., Tibshirani R., Friedman J. (2009) Unsupervised Learning. In: The Elements of Statistical Learning. Springer Series in Statistics. Springer, New York, NY. https://doi.org/10.1007/978-0-387-84858-7_14

[22] Dayan P, Sahani M, Deback G. Unsupervised learning. The MIT encyclopedia of the cognitive sciences. 1999 Oct:857-9.

[23] Wiering M, Van Otterlo M. Reinforcement learning. Adaptation, learning, and optimization. 2012 Jan;12(3).

[24] Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: A survey. Journal of artificial intelligence research. 1996 May 1;4:237-85.

[25] Ahmed N, De D, Hussain I. Internet of Things (IoT) for smart precision agriculture and farming in rural areas. IEEE Internet of Things Journal. 2018 Nov 5;5(6):4890-9.

[29] Whelan, B.M., McBratney, A.B., 2003. Definition and Interpretation of potential management zones in Australia, In: Proceedings of the 11th Australian Agronomy Conference, Geelong, Victoria, Feb. 2-6 2003.

[30] Djedouboum AC, Abba Ari AA, Gueroui AM, Mohamadou A, Aliouat Z. Big data collection in large-scale wireless sensor networks. Sensors. 2018 Dec;18(12):4474.

[31] Weyulu E, Hanada M, Kim MW. Optimizing RTS/CTS to improve throughput in ad hoc WLANs. In2017 Federated Conference on Computer Science and Information Systems (FedCSIS) 2017 Sep 3 (pp. 885-889). IEEE.

[32] Kang, Z., Yu, H., & Xiong, Q. (2013). Detection and Recovery of Coverage Holes in Wireless Sensor Networks. J. Networks, 8, 822-828.

[33] Srbinovska M, Dimcev V, Gavrovski C. Energy consumption estimation of wireless sensor networks in greenhouse crop production. InIEEE EUROCON 2017-17th international conference on smart technologies 2017 Jul 6 (pp. 870-875). IEEE.

[34] M. Srbinovska, C. Gavrovski, V. Dimcev, A. Krkoleva, V. Borozan, “Environmental parameters monitoring in precision agriculture using wireless sensor networks”, Journal of Cleaner Production, Sustainable Development of Energy, Water and Environment Systems, Volume 88, pp. 297-307, 2015

[35] Jain P, Sarangi S, Bhatt P, Pappula S. Development of an energy-efficient adaptive IoT gateway model for precision agriculture. In2018 Global Internet of Things Summit (GIoTS) 2018 Jun 4 (pp. 1-6). IEEE.

[36] Sahitya G, Balaji N, Naidu CD, Abinaya S. Designing a wireless sensor network for precision agriculture using zigbee. In2017 IEEE 7th International Advance Computing Conference (IACC) 2017 Jan 5 (pp. 287-291). IEEE.

[37] Zhang X, Andreyev A, Zumpf C, Negri MC, Guha S, Ghosh M. Thoreau: A subterranean wireless sensing network for agriculture and the environment. In2017 IEEE conference on computer communications workshops (INFOCOM WKSHPS) 2017 May 1 (pp. 78-84). IEEE.

[38] Abedin Z, Paul S, Akhter S, Siddiquee KN, Hossain MS, Andersson K. Selection of energy efficient routing protocol for irrigation enabled by wireless sensor network. In2017 IEEE 42nd Conference on Local Computer Networks Workshops (LCN Workshops) 2017 Oct 9 (pp. 75-81). IEEE.

[39] Lee, W.S. and R. Ehsani, Sensing systems for precision agriculture in Florida. Computers and Electronics in Agriculture, 2015. 112: p. 2-9.

[40] Ullah A, Ahmad J, Muhammad K, Lee MY, Kang B, Soo OB, Baik SW. A survey on precision agriculture: Technologies and challenges. InInternational conference on next generation computing, Kaohsiung, Taiwan 2017 Dec (pp. 21-24).

[41] R. K. Kodali, N. Rawat and L. Boppana, "WSN sensors for precision agriculture," 2014 IEEE REGION 10 SYMPOSIUM, Kuala Lumpur, Malaysia, 2014, pp. 651-656, doi: 10.1109/TENCONSpring.2014.6863114.

[42] Mekonnen Y, Namuduri S, Burton L, Sarwat A, Bhansali S. Machine learning techniques in wireless sensor network based precision agriculture. Journal of the Electrochemical Society. 2019 Dec 19;167(3):037522.