A report on

Cognitive Radio Networks, Topologies, Architectures and Simulations Based Analysis

Hamid Ali Hassan

May 5, 2023

Abstract

Cognitive Radio (CR) was introduced in 1999 as a concept of a device that is software programmable and re-configurable to sense, detect and eventually exploit the spectrum spaces that are not in use at all times by a user that has paid the subscription or license usage fee. Such a device or module could only be brought into reality by extending the concept of Software Defined Radios (SDRs) that are already being manufactured as re-programmable devices for similar applications in the form of reconfigurable devices. Hence, the Potential applications for CRs include Software Defined Radios (SDRs) and re-configurable and re-programmable radios for wireless networks. Communication between intelligent radio devices and associated network entities is described in it. As they learn from their experience, they can tune or change their operating parameters per the network demands. The wireless community has been enthused by such a concept that attempts to mimic human cognition and reasoning, which has sparked many research and standardization activities.
Why does someone want to use the spectrum not in use at a particular time slot? The spectrum or frequency channels, as they are called in wireless communication systems such as LTE (Long-term Evolution), and 5G (5th Generation), are costly and for the sake of not letting it be in a latent state and go to waste doing nothing most of the time, they can be utilized to maximum enhancing the efficiency and maximum possible return on the investment or capital spent for buying spectrum or frequency channels. The channel is the mini slice of the bandwidth pipe the companies buy in chunks; for example, 20 MHz for LTE bandwidth results in 20-100 KHz channels.

Federal Communication Commission (FCC) efforts have focused on allocating and starting the spectrum more efficiently using intelligent schemes. Between 2002 and 2010, the FCC issued rules and regulations allowing unlicensed devices to exploit TV white space opportunistically. This initiative has sparked an avalanche of spectrum measurement campaigns worldwide, aiming to demonstrate spectrum under-utilization.

Hence, it is essential to study how spectrum sensing works and how the channels can be sensed and used in their free time slots. The main themes of the study are the introduction of the technology, the literature review of recent, relevant and important papers, study topologies, architecture and design. However, the objective is to focus on the experimental study through simulations; the plotting of different important parameters is studied by creating multiple topologies of wireless sensor networks and CR- loT networks. The starting research is done with four topologies. Each topology is tested with different transmitters and receivers, each with its own data types, channel properties 5G, LTE, Wi-Fi, and Bluetooth Low Energy (BLE), and protocols and topologies such as star, mesh, bus, point-to-point or point-to-multi-point.
The selection of simulators is also crucial to successfully implementing the research study. Network simulators play a pivotal role in the research process. We strengthened our study by doing simulations in multiple simulators, from NS-2, NS-3, MATLAB to NetSim by Tetcos Inc. MATLAB is based on matrix calculations and extremely useful tools for plotting threshold-based energy detectors and probability-based results as well as plotting Probability distribution Function (PDF) and Cumulative Distributive Function (CDFs).
Simulating computer networks is done using NS-3, a discrete-event simulation environment. A simulator like this executes each event associated with a particular execution time. An event can generate zero, one, or more events due to being processed. When a simulation runs, it consumes events, but more can be developed (or not). A simulation will automatically end if no events are left in the queue or a special

”Stop” event occurs.
Deep learning has made its way into spectrum sensing and CR technologies. Many related algorithms finding applications in this domain are revisited, and the major cyber attacks on the CR and Internet-of-things (loT) and Wireless Sensor Networks (WSNs) are revisited. The results are presented from all these sections in the final sections and are discussed. Conclusions are drawn, and future work directions are given.

Contents

List of Tables

List of Figures

1 Executive Summary

The reason behind research on cognitive radio and spectrum sensing is spectrum scarcity, as the number of spectrum users is increasing exponentially. The advent of Internet-of-things (loT), 5G and 6G have made the spectrum a crucial asset and its usage has become highly dependent on reuse concepts, which means serving many users with the same spectrum resource to increase the efficiency of spectrum usage as well as maintain power control reduce interference to the neighbouring user in the spectrum.

The method employed is based on research based on experimentation, and simulation. The approach adopted is to create topologies of Cognitive Radio (CR) and software-defined radio (SDR) wireless sensor networks (WSN), for intermittent bandwidth-limited RF networks based on different technologies such as Long-term Evolution (LTE), Wi-Fi, Wimax, 5G, MPLS with different data rates for each case and number of nodes and plot parameters such as loss of packets, delay, interference. In the later part, the cyber security attacks such as Denial-of-service (DoS), Primary user emulation attack (PUEA) and many other cybersecurity threats that hamper the successful transmission of data between the nodes and the processing entity such as base station, cloud, or edge processor. The main parameters for plotting are delays, interference, node-link failure, jitter, and throughput; although some targets are not achieved as per objectives due to the limited duration available and the complex and broad nature of the topic from CR theory, spectrum sensing/energy detection, cybersecurity attacks and machine learning for the cyber attack prevention. The deep learning (DL) algorithms for the CR for automation of the cognitive cycle based on the training of data are also emphasized with many deep learning and machine learning (ML) techniques studied from literature for CRNs.

Satr, mesh and hybrid topologies are the most common ones and are studied. The starting research is done with a ten-node network, and the parameters such as delay, latency and throughput are recorded and studied on the terminal window. NS-3 simulations are done on Ubuntu 22.04 Linux environments and MATLAB simulations on Windows 10 Laptop.// The target included testing each topology with different transmitters and receivers, each with its own data types, each with its own channel properties 5G, Longterm Evolution (LTE), Wi-Fi (Wireless Fidelity), Bluetooth Low Energy (BLE) and protocols and topologies such as star, mesh, bus, point-to-point or point-to-multipoint.
Video transmission and reception are, however, not very common in bandwidth-limited intermittent networks. The WSN or loT or CR-loT networks are also power/bandwidth limited, which deters video messages since video messages require large bandwidth pipes.

After the architecture and topologies, the literature shifted towards deep learning implementations, energy detection methods (co-operative and non-cooperative), deep learning methods for spectrum sensing and CR, and the cybersecurity risks and vulnerabilities in future CR systems. The main attacks that we studied are Dos, primary user emulation attacks, and spread spectrum data falsification attacks.

The main parameters of interest for simulations are false alarm probability for malicious users, CDF for secondary and primary users, transmission delay from the transmitter to receiver, latency, throughput, power control and average transmit power etc.
Finally, the simulation results are presented, and the discussion, conclusions and future directions are provided.

2 Introduction

2.1 LITERATURE REVIEW

Government regulations and licensing fees for radio spectrum are high, and telecommunication businesses pay hefty amounts for the frequencies. The most significant cause of spectrum scarcity, as per the Federal Communications Commission (FCC), is inefficient primary user spectrum utilization. Developing technology such as co-operative spectrum sensing (CSS) and cognitive radio (CR) possesses the capability to alleviate the issue of spectrum shortage. By utilizing opportunistic spectrum sharing in CR, it improves the utilization of the spectrum. The idea behind the concept is that the secondary user (SU) uses the frequency channels or time slots not used by the primary or licensed user at one particular instance (this window of opportunity is used as white space). CR performs opportunistically that does not interfere with the primary user (PU) in a particular time slot. [10].

CR was introduced in 1999 by Dr. J. Mitola III, with potential applications such as Software Defined Radios (SDRs) and re-configurable radios over wireless networks. Communication between intelligent radio devices and associated network entities is described in it. As they learn from their experience, they can tune or change their operating parameters per the network demands. The wireless community has been enthused by such a concept that attempts to mimic human cognition and reasoning, which has sparked many research and standardization activities [27].

FCC efforts have focused on allocating and starting the spectrum more efficiently using intelligent schemes. Between 2002 and 2010, the FCC issued rules and regulations allowing unlicensed devices to exploit TV white space opportunistically, such as [54, 88] MHz, [174, 216] MHz, and [470, 806] MHz, where each TV channel has a width of 6 MHz. This initiative has sparked an avalanche of spectrum measurement campaigns worldwide, aiming to demonstrate spectrum under-utilization [46].

The ”Cognitive Radio” concept defines a radio that changes its transmitter parameters in response to the environment in which it operates. In light of this definition, it is possible to define the CR in terms of two main characteristics:

2.2 The Cognitive Capability:

The cognition of a radio technology refers to its ability to sense or collect information from its radio environment. For this ability to be actualized, more sophisticated techniques are needed to capture the temporal and spatial variations in the radio environment and avoid interference with other users. Spectrum utilization can be identified at a specific time or location by looking at the unused portions of the spectrum. It is, therefore, possible to select the appropriate spectrum and radio transmission/reception parameters based on these obtained statistics. Spectrum awareness is provided through cognitive abilities, while re-configurability allows for dynamic firmware programming dependent on the radio environment. More specifically, the hardware design of the CR allows it to enable the use of different transmission access technologies and several frequencies for transmission and reception [31].

Cognitive radio networks (CRNs) divide their users into two classes: the main users are referred to as primary users (PUs), and to coexist with them; CRs also provide secondary users (SUs) that share the same spectrum resources with PUs without causing significant interference. Innovative/intelligent policies enable SUs to intelligently utilize spectrum resources in CRNs by automatically sensing licensed channels authorized to PUs. Based on intelligent, opportunistic access, CR technology integrates traditional wireless networks with new technologies. CRs rely on the composition of wireless networking protocols layer by layer to form a robust communication network. On the other hand, they extend frequency resource availability by using opportunistic spectrum sharing (OSS) [12].

2.3 Selection of Network Simulator

The selection of simulators is also crucial to successfully implementing the research study. Network simulators play a pivotal role in the research process. Various research studies have examined the efficiency of multiple methods and protocols for routing in multiple simulators with varying network parameters. A commonly adopted preferred option is NS-2 and NS-3. Some other simulators include GloMoSim/QualNet, OMNeT++, TOSSIM, OPNET Modeler Wireless Suite, MiXiM, Castalia, INET framework, NesCT, Avrora, NS-2, J-SIM, ATEMY, Emstar, SENS, SENSE, and SHAWN [6].
The nodes in a WSN can be simulated with MATLAB scripts, and their relevant energy consumption, network lifetime, node coverage, and clustering-based nodes (Cluster Heads and Sinks) can be plotted. MATLAB is suitable for CRN simulations but is only limited to the energy detection part of spectrum sensing that stems from information/detection theory and coding and the machine learning of detection and the cyberattacks part. A good example of MATLAB for detection in SS is given in this reference [17]. Another excellent example of the MATALB for CR implementation is described as a step towards future wireless technology in this study[38].
This study mainly studies CRNs by simulating different topologies in NS-3 and/or OMNET. The primary method for studying CRNs is analytical, or simulation methods since the protocols and techniques that are not adequately tested and verified can't be launched at a larger scale and deployed in real-world scenarios. A comparison analysis of NS-2, ns-3. OMNET++ and GloMoSiM have been performed for four different simulation environments based on the comparison of algorithms/protocols for routing, such as Ad-hoc On-demand Distance Vector (AODV). The researchers simulated the AODV routing protocol in four different simulation environments that are NS-2 and NS-3. OMNET++ and GloMoSiM. The performance comparison is based on the parameters such as the number of sensor nodes vs Memory usage, computational time and CPU utilization. The NS-3 simulator is among the fastest in computation time among the simulators selected [23].
Another study compares the simulation software NS-2, NS-3, OMNET++, JiST, and SimPY based on parameters such as simulation run-time vs network size, end-to-end packet loss, simulation run-time vs drop probability, memory usage vs network size, memory usage vs drop probability. The JiST simulator has proven to be the fastest in this study based on simulations. There are, however, some simulation scenarios where its applicability may be limited by its total exhaustive memory consumption. Based on our performance comparison, NS-3 showed the best performance overall. Despite being surpassed by JiST in simulation run-time, it has a low computational and memory requirement. NS-3 is still in its infancy, and only a few ready-made simulation models exist. The rich collection of models for NS-2 still needs to be ported from NS2 to NS-3, so OMNeT++ can be considered, even though its performance lags behind NS-3 and JiST [43]. Two scenarios are considered for the performance comparison of different network simulators, with 50 nodes in one design and 100 in the other. The simulation is performed with the same parameters on different simulators such as OMNet++, NS-2, NS-3, and MATLAB/Simulink, and the results were surprisingly faster for MATLAB/Simulink. MATLAB exhibited a remarkable decrease of 94.11%, 93.10%, and 92% in simulation run-time when competing with other WSN simulators such as NS-2, NS-3, and OMNET++, [32]. The parameters are given in Table 1, and the results are displayed in Figure 1.

Table 1: Simulator Performance Comparison for WSNs.

Simulator	Scenario 1, 50 nodes (ms)	Scenario 1, 100 nodes
NS-2	3.4 m	3.6 ms
NS-3	2.9 ms	2.6 ms
OMNet++	2.5 ms	2.7 ms
MATLAB	0.24 ms	0.7 ms

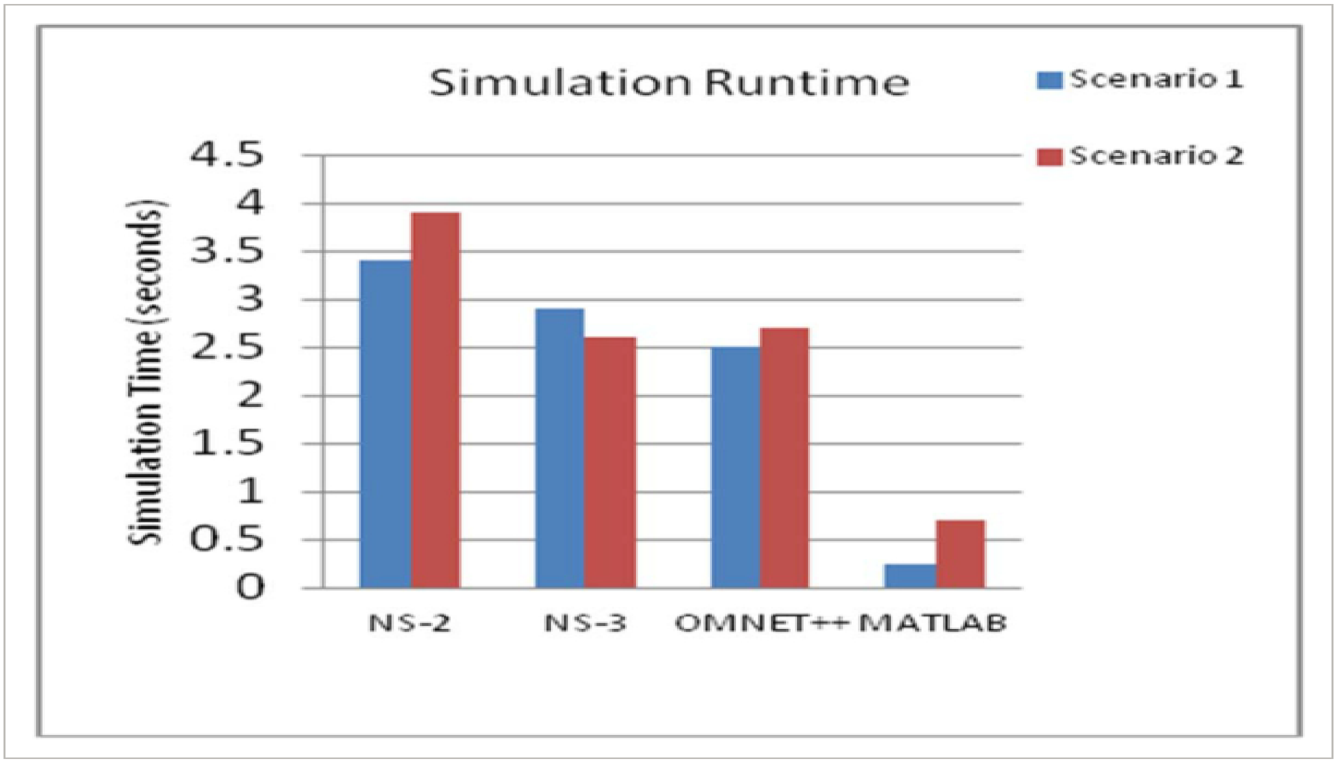

Figure 1: Run-Time Performance Comparison Of Different Simulators

The NS-3 simulator can be helpful for many application scenarios such as LTE, Wi-Fi, WiMAX, 5G, and it offers integration to end-to-end TCP/IP protocol stack. Deep learning has penetrated many application areas. OpenAI gym offers a dedicated toolkit for NS-3, named NS3-gym for developing and deploying reinforcement learning (RL) algorithms. Many agents can be developed in OpenAI gym, such as playing video games such as ping pong, robotics and hanging pendulum problems. There are many advantages to OpenAI-gym for RL, like the availability of Tensor flow and Scikit-Learn and the easy coding of agents through python. There are scalability benefits, low complexity conversion and integration of codes into OpenGym from NS-3, quick prototyping, and easier troubleshooting maintenance of Open-gym codes. This study examines the performance of RL-agent learning in a CR example in a multichannel wireless network setting, such as 802.11 networks, with outside interference, [15].

CR technologies can exploit and use both licensed ad unlicensed bands. In licensed bands, the CRs use the spectrum holes while the primary user paying for the license is always given priority. In unlicensed bands, all users have the same preference. The ISM bands at 2.4 GHz and 5 GHz are examples of unlicensed bands.

When choosing architecture/topology for wireless sensor networks (WSNs), the main objective is reducing interference caused by simultaneous transmission. Exploiting 802.15 is a preferred route to achieve

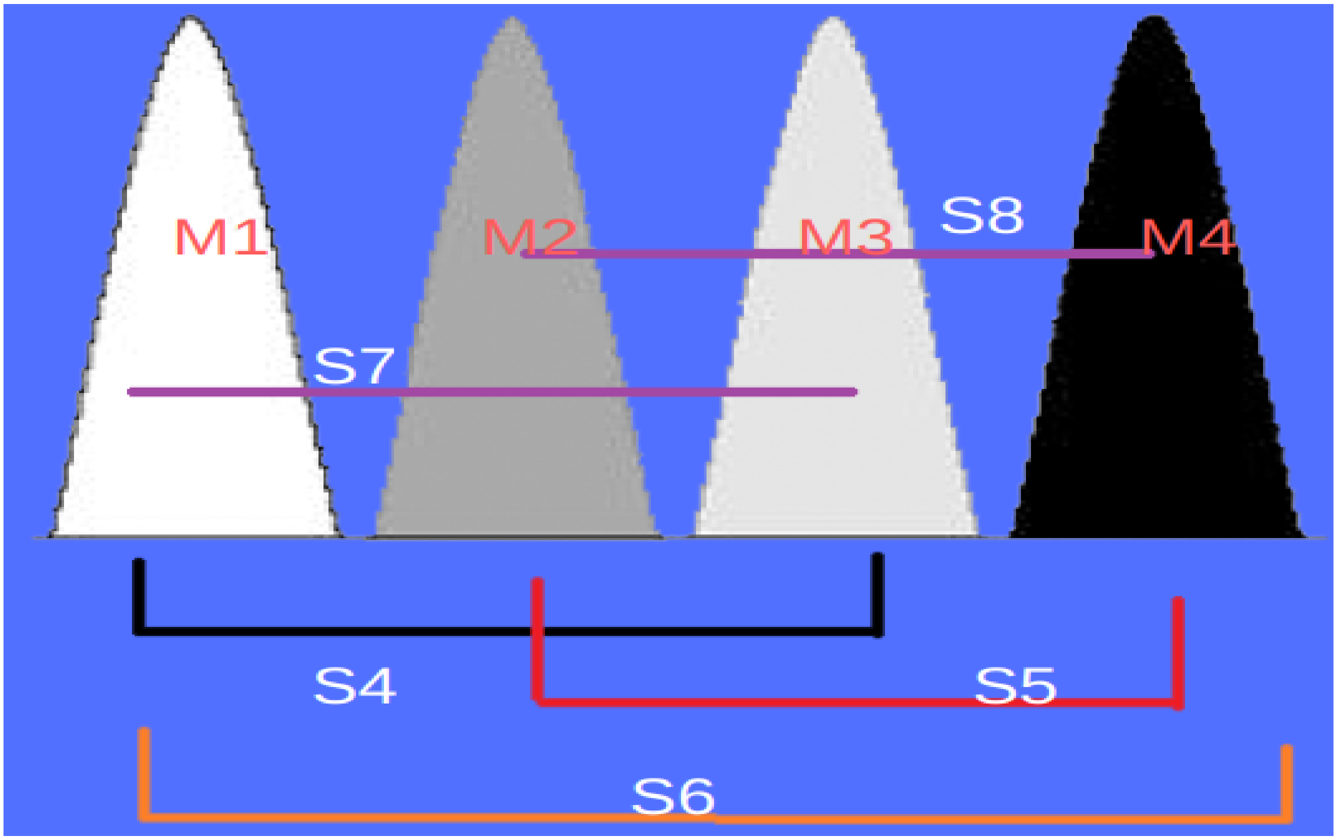

Figure 2: Four channel-based allocations in case of CSS

this goal, while many tree-based topologies have been implemented in different research attempts, [44].

2.4 Defining Cognitive radio

As part of CR, unused spectrum is repurposed opportunistically to maximize underutilized spectrum resources. A CR system includes primary users, who are incumbent licensees, and secondary users, who seek to make use of the spectrum opportunistically when the primary users are idle. For CRs to determine if the spectrum is available, they must sense it, and the reinforcement learning type system can take the appropriate action.
The cognition cycle in CRN consists of four main phases, which can be approximated as sensing, analysis, reasoning and adaptation. The cognitive engine is hosted as part of the Cognitive Control Server, including the spectrum sensing and monitoring algorithms, learning and testing functions, and a database to store the knowledge and case-based decision-making. CCS (Cognitive Control Server) is responsible for brain functions. Based on the assumption that we have four channels, we have a probability of two to the power four for the primary user to exist [those probabilities are assigned a state ID]. Hence the CSS determines what state is available and allocates it to the secondary user for use while it is still unused. Each state has a different modulation scheme, data rate, transmit power, etc. This situation is illustrated in Figure 4.
According to CR technology, CRNs serve the following purposes:

- The spectrum can sense or determine the spectrum band and detect the presence of PUs.
- Aspects related to spectral efficiency, such as selecting the most appropriate frequency spectrum sensing system to meet the users' needs or communication requirements.
- Coordination of access to this broadcaster with other customers through spectrum sharing.

If a PU is observed or detected, spectrum analysis may be conducted, or the broadcaster is vacated. CR distinguishes itself from conventional radio in three ways.

- Cognitive: As part of its operational and physical surroundings, CR is aware of its location.
- Reconfiguration: CR can change its parameters dynamically and independently based on this cognitive understanding.
- Learning:. When new situations arise, CR can apply what it has learned to new setups, [19].

The main topic under research is spectrum sensing using CR and SDR. It is essential to consider the various techniques for spectrum using state–of–the art literature and classify them based on some common properties. The issues and challenges in implementing these technologies are also discussed, [33].

3 Theory, Topologies, Architectures, and Scenarios.

Keywords: Internet of Things (IoT), Cognitive Radio Networks (CRNs), Artificial Intelligence, Machine Learning.
The cognitive radio network (CRN) comprises two critical systems: the core and subsidiary systems. An example of a core radio base station is the central system that owns the licensed band due to the split bandwidth between the core and secondary networks. The CR system is composed of users and a ground station.
This software-defined radio framework enables CR to understand and adapt to the transmission medium to become a situational smart radio independent of human customization.

3.1 The architecture for CR:

Regarding architecture, CRNs can be classified into three branches: infrastructure, Adhoc, and mesh architecture. Infrastructure is centralized, based on Access points (AP) and Mobile stations (MS) communicating with each other within the range of AP in a one-hop manner. In Ad-hoc, there is no centralized AP. MS is a mobile station, which are CRs in this study, communicate with each other using licensed/unlicensed bands. Mesh is a hybrid of both ad-hoc and infrastructure. Hence, the nodes communicate with each other directly, or through centralized AP, [26].

3.2 The cognitive cycle:

Founded on a series of actions encompassing sensing, data fusion and information aggregation, and deciding process, CR networks could be described as intelligent components that help allocate scarce radio frequency spectrum to demanding users/consumers, [21].
The cognitive cycle is an operation majorly consisting of three operations. Figure 3 illustrates a typical cognitive or mental cycle. Figure 2-1 shows that the beginning of the cognitive cycle is given by a systematic chain of sensing acts, leading to the aggregation of perceived information and, ultimately, to final decision-making based on the sensed information[33]. The environment functions to provide the initial stimuli. To extract the contextual cues that are readily available for carrying out the tasks that have been assigned to it, CR analyses these inputs. One might consider GPS coordinates with light and temperature to establish whether it is inside or outside a building. This kind of processing occurs in the cognition cycle's observation phase. Content from incoming and outgoing messages, including those supplied by the user, is processed, [28].

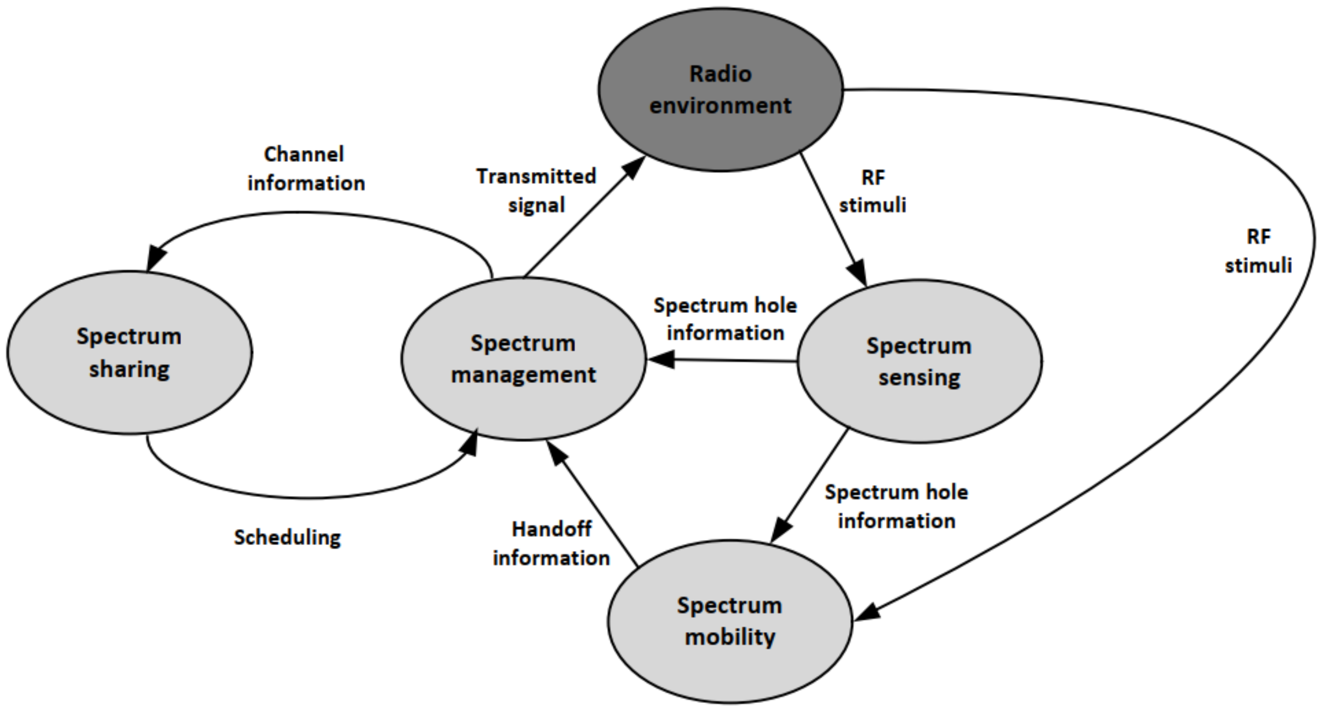

Figure 3: Spectrum Sensing Cycle in Cognitive Radio Theory[18].

The main types of methods are cooperative and Non-cooperative sensing. In the cooperative mode of sensing, the SUs applying for access to the spectrum do so without consulting with other SUs about their objectives. Only through efficient and continuous sensing of channels can this vacancy be detected. White space or spectrum void are different names for the spectrum opportunity created, termed Television White Space (TVWS) [3].

3.3 Spectrum Sensing

Spectrum sensing is finding open frequency ranges, often known as spectrum holes. Because they are accessible in time or space, spectrum holes may exist. Through the spectrum sensing technique, CRs can gauge and understand the environment of operation and form an awareness of the environment in terms of frequency channels available and the respective interference constraints. When the PU is not using a band at a particular moment or time slot, secondary users can take advantage of the spectrum, i.e., spectrum opportunities exist. This makes spectrum sensing possible in various time, frequency, and spatial domains. The development of beam-forming technology has allowed multiple users in the exact geographic location to use the same channel and frequency simultaneously. In this scenario, extra frequency channel vacant slots are generated for secondary users in the particular slots where the primary user is not transmitting, and spectrum sensing must also consider angles of arrival (AoAs), [42].
There are two main classes for the detection of energy of the vacant channel, one of which is non-cooperative, also termed as transmitter detection, and the other is co-operative energy detection method, [7]. Approaches for detecting energy in transmitters/non-cooperative methods rely on CR users' close-by observations of signals sent from a primary system. Most transmitter or non-co-operative detection methods are built on the idea that the apparatus performing cognition is unaware of the primary transmit-

ter's position. Therefore, to accomplish spectrum sensing, cognitive users should only rely on picking up weak and limited local information to perform sensing the available to-use frequency channels. Moreover, the cognition equipment only has limited local information of the environment in which it searches for vacant frequency channels, [37].
Consequently, it is improbable not to receive malicious interference signals from licensed users altogether. Transmitter/Non-cooperative always has the problem of a remote, hidden terminal that it can't avoid altogether. Three non-cooperative/transmitter detection methods are popular and commonly available in the literature. These are matched filter detection, energy detection, and cyclo-stationary feature detection, [48]. Figure 4 illustrated the different spectrum sensing methods from cooperative and non-cooperative detection classes.

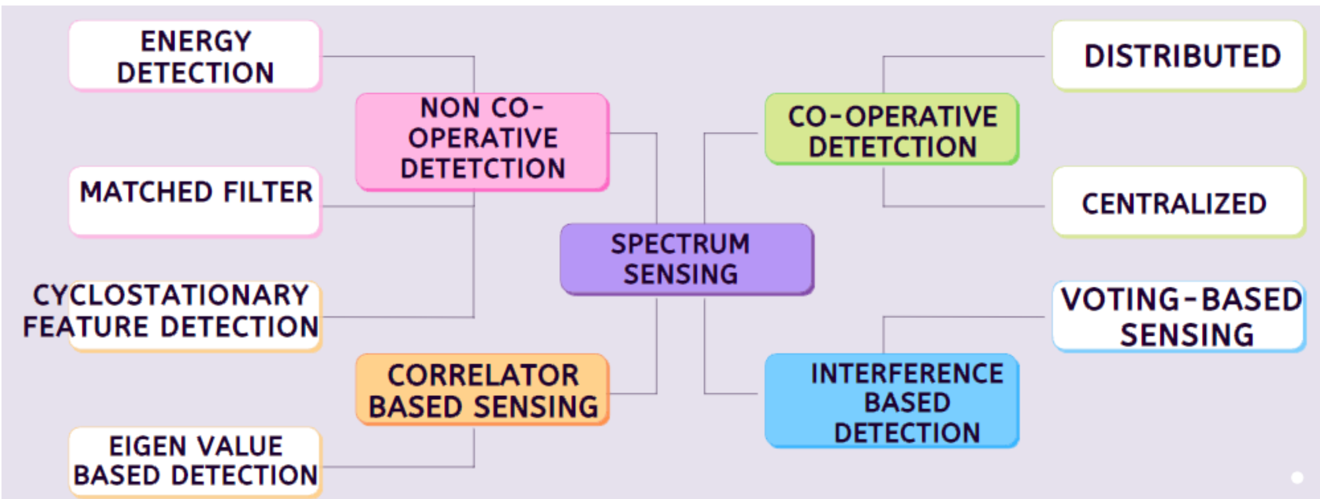

Figure 4: Spectrum Sensing and Energy Detection Methods

Due to the shadowing phenomena, which is highly common in urban/indoor situations, a cognitive user (CU) may have a strong line of sight with a primary receiver but may not be capable enough to detect the existence of a primary transmitter (a problem known as hidden terminal). To lessen this issue, cooperative detection procedures are used. For more precise primary transmitter detection, cooperative detection refers to spectrum sensing techniques that let many CRs share their local sensing data. Either centralised or decentralised implementations of cooperative detection are possible. In the centralised approach, a central unit gathers sensing data from cognitive devices, identifies the available spectrum bands, and broadcasts this data to other CRs. There is no central node in a distributed strategy, and the sensing data is dispersed across the cognitive devices. While centralised detection is more accurate and can reduce multi-path fading and shadowing effects, distributed detection is simpler to install and does not need back-end infrastructure support. [4]. Figure 5 illustrates how the PUs use the allocated spectrum or frequency channels and the situation after SUs occupy the empty frequency channels or slots.

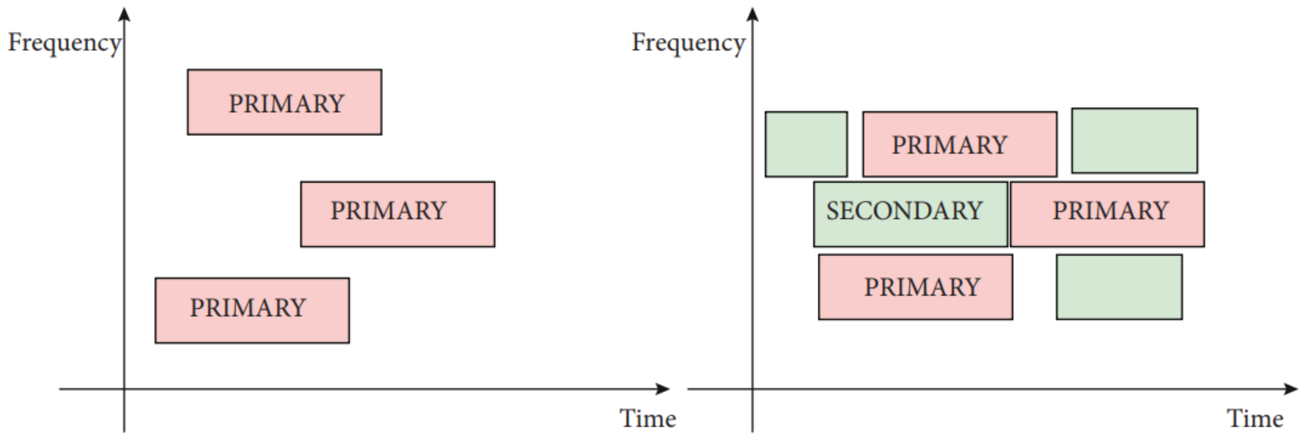

Figure 5: Primary User (PU) and Secondary User (SU) Before and After Empty Spectrum Spaces Allocation.

3.4 Energy Detection

This is the most prevalent and widespread method for empty frequency channel scavenging, otherwise known as SS. It is based on the detection of the samples received signal. The main attraction of their method is that it does not need prior information on the primary signal. As well as ensuring a certain level of primary-user protection, a good detector should ensure a greater detection probability and a small probability value for the false alarm. Various approaches have been proposed for improving energy-detector-based spectrum sensing efficiency, [41]. The disadvantage of this system is the high false-alarm rate caused by noise uncertainty. Also, the low-SNR cases are unreliable as they can't distinguish between primary and secondary users.

3.5 Coherent Detection

3.5.1 The matched filter detector (MFD)

The matched filter, a coherent detection method, exploits a correlator that is matched to the signal energy of relevance or, to some degree, correlation of it. Common examples are pilot bits and training bits. This method involves comparing the received signal to the PU signal to determine whether PU is there or not. Matching filter detection is the most successful technique, assuming Gaussian noise is present [49]. However, the cognitive user must be completely aligned and have information about the PU to apply the matching filter detection, which is often not feasible, especially at low SNRs. The matching filter approach locates a signal by calculating the correlation coefficient of the signal received and a signal already available to be duplicated. It requires complete signal information from the primary user, including the operating frequency, bandwidth, a form of modulation, the order, waveform properties, and structure packet, to be the optimum detection tool. Additionally, detection performance will worsen if erroneous data is used for matched filtering, [20].

3.5.2 Cyclo-stationary Feature Detection (CFD)

Identifying primary signals based on their deterministic or statistical characteristics is the foundation of feature detection. Feature detection can identify between signals with various features based on extracted signal features. In general, feature detection is more computationally complex than matched filtering or energy detection. Cyclo-stationarity-based feature detectors are a significant subclass of feature detectors more resistant to noise uncertainty than energy-based feature detectors since noise is often not cyclo-stationary. However, synchronization faults can significantly impact cyclo-stationarity-based detection, leading to carrier frequency and sample clock frequency discrepancies. Because it can distinguish the primary signal from interference and noise, the cyclo-stationary feature detection method utilized in CR is a very alluring frequency channel sensing scheme. By exploiting the stochastic variable, such as mean and auto-correlation of the signal being received, the cyclo-stationary feature detector detects these cyclically occurring parameters of signals, which are not random. The received signal is connected to the primary user if the mean and auto-correlation vary periodically; otherwise, it is noise, which is not periodic. Because of this, cyclo-stationary feature detectors can distinguish between the primary user signal and noise even in conditions with very low SNR. It is challenging to maintain synchronization in low-SNR circumstances for this form of detection, though. Fig. 6 depicts a schematic of cyclo-stationary feature detection, [1].

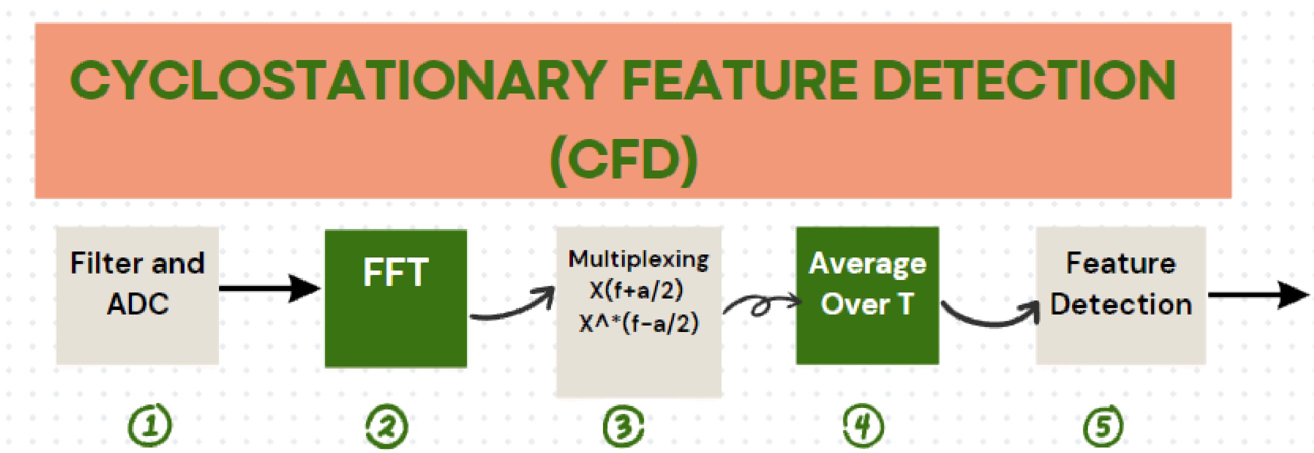

Figure 6: Cyclo-stationary Feature Detection.

Regarding computation power, the cyclo-stationary feature detector exhibits complex behaviour, whereas the MFD assumes that the CR has complete knowledge of the original signal. Several machine learning (ML) algorithms have been researched and developed to address these issues, such as reinforcement learning, game theory, support vector machines (SVMs), K- nearest neighbour (KNNs), ANN, CNN, etc.

Criteria	Energy detection	Matched filter detection	Cyclostationary feature detection	Covariance-based detection	Waveform-based detection
Detection accuracy	Not good at low SNRs	Optimal for all SNRs	Excellent performance at all SNRs	Moderate performance at all SNRs	Superior performance at all SNRs
Complexity	Low	High	High	High	Intermediate
Information of PU required	No	Yes	Yes	No	Yes
Robust for noise uncertainty	No	Yes	Yes	Yes	Yes
Sensing time for better performance	Low	Low	High	High	Low

Table 2: Comparison of spectrum sensing techniques from co-operative spectrum sensing

A comparison of different energy detection techniques from co-operative spectrum sensing is presented in Table 2 based on accuracy, complexity, noise suppression, information requirement of PU user and sensing time metrics, [22].

3.6 Challenges and Issues in Spectrum Sensing.

When addressing the spectrum sensing problem in CRNs, several sources of uncertainty, including channel uncertainty, noise uncertainty, sensing interference limit, etc., must be considered. In the case of wireless networks, channel fading or shadowing can cause inaccuracies in received signal strength, leading users to disbelieve that the primary system is outside of the secondary user's interference range even though the direct signal may be suffering from a deep fade or being heavily obscured by obstacles. The same challenges lie in aggregate interference and noise uncertainty, [36]. Some common challenges in Spectrum Sensing are illustrated in Figure7.

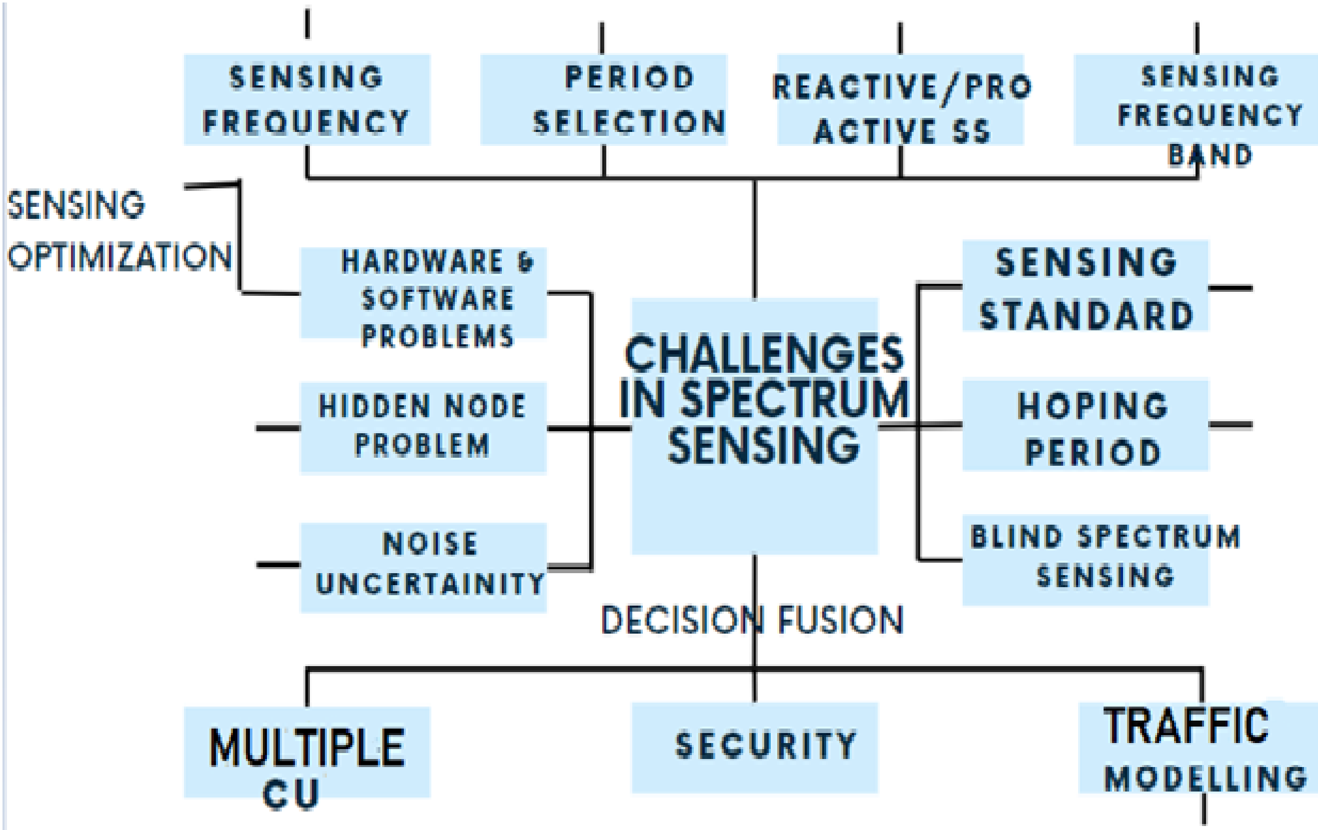

Figure 7: Challenges and Issues in Spectrum Sensing.

4 Deep learning in Cognitive Radio Networks

Keywords: Deep learning (IDL), Reinforcement Learning (RL), Support Vector Machines (SVM). Long sort-Term Memory (LSTM), Machine Learning (ML), Random Forest (RF).
Deep learning has penetrated every application area, and modern communication systems are no exception. This chapter studies machine learning and deep learning algorithms applicable to CRNs. The model training procedure is presented from the start to the end of the model training chain. The algorithms from deep learning and machine learning suitable for CRN applications in Software Defined Radio are reviewed.

4.1 Introduction to Machine Learning in CRN Theory:

Considering the vastness of CRN theory, primary attention has been given to spectrum sensing, dynamic spectrum access (DSA) and small topics such as cooperative sensing, spectrum prediction and auctioning. While interference control must avoid harmful interference with a primary user, SS's main challenge is detecting white spaces (PU). Multiple CRs on a PU can interfere with one another; hence distributed Q-learning can reduce this. DSA is essential to optimize the spectrum utilization once these white spaces have been identified. Some standard spectrum management methods include game theory and reinforcement learning.

4.2 Open AI Gym and Reinforcement Learning:

A toolset for reinforcement learning (RL) research is called OpenAI Gym. It contains many well-known problems that reveal a standard interface, enabling direct comparison of the performance outcomes of various RL algorithms. The de-facto standard for academic and commercial research into networking protocols and communications technology has long been the NS-3 network simulation program. Hundreds of models and modules have been built and added to the NS-3 code base, and numerous scientific publications have been published presenting results achieved using NS-3. The use of machine learning technologies like RL is an important current trend in network research. The network simulator NS-3 lacks an RL framework integration, such as OpenAI Gym. The NS3-gym framework is presented in this publication. We first go over the choices made when designing the software. Two examples that use NS3-gym are shown as the second step. Our open-source software package is available to the public and easily extensible, [14].
Some advanced deep learning algorithms from deep learning and machine learning are tabulated in Table 1.

Some Advanced Deep Learning/Machine Learning Algorithms
OpenAI Gym, Reinforcement learning
CNN, ANN, Q-Learning
Multi-Agent Q-Learning, Multi-Agent Deep Q-Network
Unsupervised Learning (GMM, K Nearest Neighbor (KNN), K-Means)
Directed Acyclic Graph Support Vector Machine (DAGSVM), Backpropagation (BP)

Table 1: Some of Advanced Deep Learning Methods for SS and CRNs

4.3 Common Machine learning Techniques:

The most common machine learning techniques fall in the category of supervised learning and unsupervised learning. The former uses labelled data, while the latter exploits unlabeled data for processing. Clustering is an example of unsupervised learning, while regression and classification are examples of supervised learning.

4.3.1 Supervised learning

If the data is labelled, it can be divided into various classes based on the labels. Applying supervised learning strategies like regression and classification is the equivalent of this. In this particular scenario, regression and classification are examined. Regression Using an equation, the linear regression technique can express the relationship between two variables.

$$h_\theta(x) = \theta_0 + \theta_1(x) \tag{1}$$

where $h_\theta(x)$ = hypothesis and $\theta_i = value\ of\ paremeters\ to\ be\ chosen\ during\ the\ training\ process.$

$$y = ax+b \tag{2}$$

Table 3: Linear Regression and a Linear Equation in Slope-Intercept Form

The output of the continuous-time real-value model can be predicted using the second equation in Table 3. A training set of data is fed through a learning algorithm (a common example of which is gradient descent), and a hypothesis function is created using an example of the size of a house to be predicted. The model will provide a forecast or anticipated price of the home if the size of the house is run through this hypothesis. From the perspective of algorithm implementation, some significant regressions kinds include:

1. Polynomial Regression
2. Logistic regression
3. Lasso Regression
4. Ridge regression

Classification Classification is the best strategy for predicting a discrete value; regression only predicts continuous real-valued output. The way that data is categorized is by grouping it according to how similar its features are. The finest example of classifications is provided by support vector machines, which are thoroughly discussed for their application in CRNs in the algorithm's examples section.

4.3.2 Unsupervised learning

A common example of unsupervised learning is the clustering scheme which uses unlabeled data and processes it as per the techniques demonstrated by the clustering mechanism. Clustering The idea of clustering refers to comparable grouping characteristics or object traits. While this point is being debated, it could be argued that classification, which involves assigning objects with similar characteristics to one class, accomplishes the same thing. However, classification is supervised learning, whereas clustering is an unsupervised learning technique, meaning that the developed data is unlabeled and clusters are formed based on similarity. K-Means and DBSCAN are two of the most used clustering methods. Stuart Lloyd, a bell laboratories employee, proposed K-Means in 1957 as a technique for pulse-code modulation, although the business didn't release the idea until 1982. The study was titled ”Least square quantization in PCM”. Clustering can be applied to segmentation, semi-supervised learning, pre-processing, and anomaly detection using Gaussian mixtures, [16].

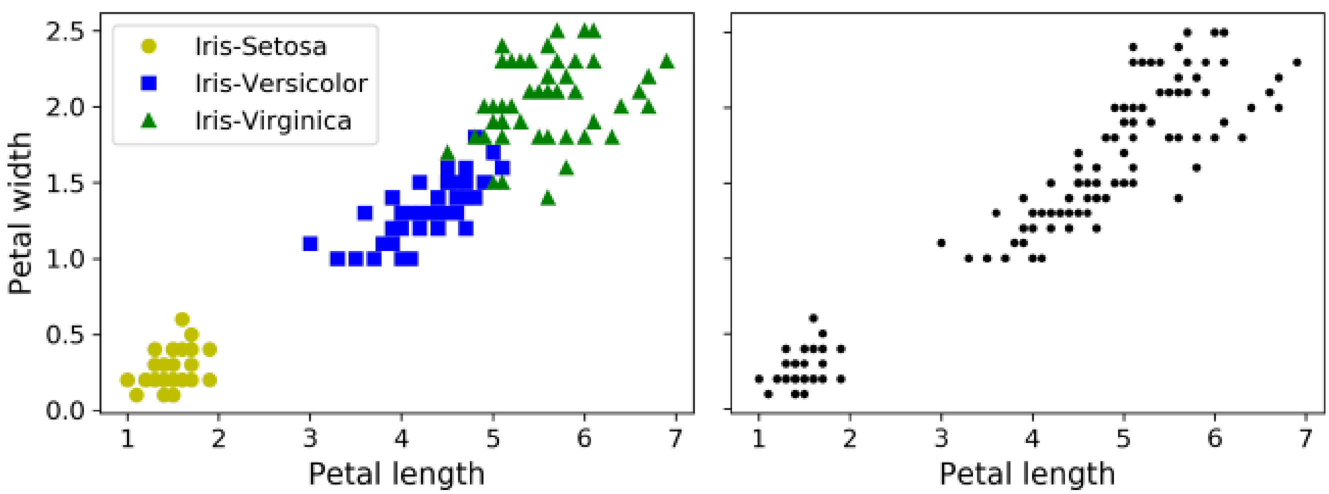

Figure 1: Classification (left) vs Clustering (right)[32].

4.3.3 Anomaly Detection

An excellent illustration of anomaly detection is the measurement of characteristics like heat generated and vibration intensity in an aeroplane engine. Anomalies must be found when a new aircraft engine is put into service to determine if it must undergo additional testing. If the features of the outliers are less than or equal to a predetermined threshold, they are marked as anomalous. Typical uses include fraud detection, industrial processes like proactive maintenance, monitoring of computers, CPU load, and network traffic in data centres. According to [5], other important anomaly detection algorithms are Fast-MCD, isolation forest, local outlier factor (LOF), and one-class SVM.

4.4 Algorithms Examples From Supervised/Unsupervised Learning Machine Learning Techniques.

Deep learning and machine learning provide various algorithms suited for wireless communications applications, including CRNs. In this study, light is shed on the algorithms that have CRN and SDR applications and are specifically related to communication systems in modern communication systems.

4.4.1 SVM (Support Vector Machines)

A classification method known as support vector machines (SVM) uses the concept of hyperplanes as support vectors. In n-dimensional space, a hyperplane is an n-1 subspace. One dimension can be added to a two-dimensional plane, and two dimensions can be added to a three-dimensional subspace using a two-dimensional hyperplane. This concept is similar to splitting two classes using a simple linear line. The data are referred described as being ”linearly separable”. The dividing line is known as a decision boundary. SVM, which can distinguish between linear and non-linear relationships and even do outlier detection and regression, is undoubtedly the most effective machine learning technique. Because most data gathered from sensors or other sources is not linearly separable, SVM contains the concept of kernels like polynomials, Gaussian RBF, [5]. Performance of SVM algorithm FPR and TPR are used to describe the effectiveness of the proposed SVM algorithm-based MUs identification in CR-IoT networks. The following values can be used to evaluate these parameters:

- True Positive (TP): Successful identification of a typical CR-IoT user.
- True Negative (TN): Succeeded in correctly identifying a rogue CR-IoT user.
- False Negative (FN): When a trustworthy CR-IoT user is mistakenly identified as hostile.
- False Positive (FP): A hostile CR-IoT user was mistakenly identified as legitimate.

Now, the (True Positive Rate) TPR and (False Positive Rate) FPR are determined using the formulas as given in Table 4 below:

$TPR = 1 - \frac{FN}{TP + FN}$	(3)
$FPR = 1 - \frac{TN}{FP+TN}$	(4)

Table 4: TPR and FPR formulas

Therefore, the TPR and FPR must be constrained to 1.0 and 0.0 for effective classifiers, respectively, [47].

4.4.2 Random Forest

The random forest algorithm is a member of the group of machine learning techniques known as ensemble methods. According to theory, integrating or aggregating the results will be preferable if the data includes many forecasts. That will increase precision. The premise behind ensemble approaches is this. There are two categories: boosting-based strategies like Adaboost and XGboost and tree-based methods like a random forest. Combining multiple classifiers' results, the bagging idea is employed, [8]. Another powerful technique with communication system uses, such as CRN or SDR, is random forest.

4.4.3 Recurrent Neural Networks:

Recurrent neural networks (RNNs) are a type of network that uses previous results as inputs to modify results in the present. To store and use information about the previous output state for subsequent outputs is analogous to having a network memory. The configuration of RNNs can be one-to-one, one-to-many, many-to-one, or many-to-many. Tanh, Relu, and sigmoid are the most common activation functions used by RNNs. One example of the SS in CR Using CNN-RNN and Transfer Learning is provided in [34].
A few of the many applications for these technologies include language translation, also known as machine translation (incorporated in Google Translate), natural language processing (NLP, integrated as Siri, Voice search), speech recognition, sentiment classification, image captioning, and music production. This study explores RNN for PUE Attack Detection in practical CRN models, [11].
Recurrent neural networks can be divided into two basic categories: LSTM (Long Short-Term Memory) and gated-recurrent units (GRUs). LSTM is briefly described in the next section.

4.4.4 LSTM (Long Short-Term Memory)

This artificial recurrent neural network technology is derived from deep learning and natural language processing (NLP). This differs from RNN and conventional CNN (Convolutional neural network), which lack a feedback connection. Both speech and image data (single point data) can be processed by LSTM (data sequence). Three gated components—an input, an output, an exit, and a cell—make up an LSTM cell. Gates are used to enter and leave information from cells, which contain information about numbers. The vanishing gradient problem that arises when LSTM resolves training traditional RNNs. LSTM Based frequency channels sensing method for CR exploiting preliminary activity parameters is studied in [35]. Another Deep Learning-Based SS attempt in CR has been conducted using A CNN-LSTM Approach [45]. A taxonomy of the deep learning and Machine learning techniques is presented in Figure 2. An extensive survey of Machine learning algorithms covering genetic algorithms, self-organizing maps, Q learning, K nearest neighbour, deep reinforcement learning, hidden markov models, Artificial neural networks, Long short term memory and game theory/decision theory is presented in this study [40].

4.5 Classification of Attacks on CR-loT Networks (CR-loT)

There are many attacks with vulnerabilities at every layer of CRN architecture.

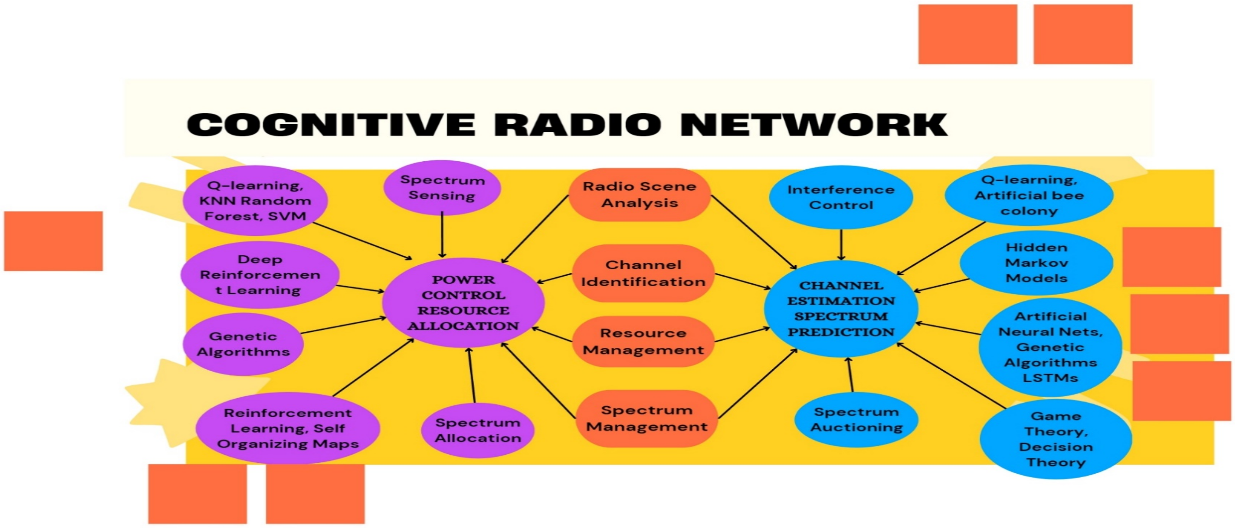

Figure 2: Machine Learning Techniques in CRN Theory and Main Spectrum Sensing Functions..

4.5.1 Node failure:

The intrinsic constraints of loT battery power cause nodes in the Internet of Things networks to fail. Sometimes nodes are abandoned in hostile environments where they perish, and battery replacement is not an option. The network lifespan is directly correlated with this. Nodes in a network can malfunction in CRNs for various reasons, including faulty node hardware, a broken network, software flaws, etc. The first node failure and the period of half-node loss will immediately affect the network lifetime. To measure the dependability of a MANET network and its influence on network performance satisfactorily, nodes can fail after a given length of time.

4.5.2 Denial-of-Service (DoS)

DoS means denying the resources to the licensed user by the interferer or the unauthorized user. The attacker can use the power of the transmitters to disrupt regular communication thanks to advances in hardware technology. The communication may be impeded by adding a noise spectrum to normal communication signals. Because of this, users' access to wireless base station equipment would be denied due to a lack of resources. Interfering with information will have a significant social impact. For instance, the lawbreaker who installed a VSAT terminal using high power to disrupt satellite service was responsible for the incident of Xin's communication satellite interference that occurred in 2001 [25].

4.6 Create and Detect A Primary User Emulation (PUEA) Attack In Cognitive Radio Networks.

A promising technology, CR, allows unlicensed users equipped with CRs to co-habit with licensed users in licensed frequency bands without interfering with the communications of the licensed users. Spectrum Sensing is the most essential phenomenon of CR, with continuous research and development on the operation and development aspects of the technique and its parent technology. The three main pillars of security are confidentiality, availability and integrity. In a hostile environment, an attacker may alter a CR's air interface to imitate the characteristics of a primary user signal, leading genuine secondary users to think the attacker is a prime user mistakenly. We designate this attack as a primary user emulation (PUE) attack. An attacker impersonates a primary user or incumbent by sending a signal with traits of a primary signal or responding to a true one. The CRN cannot use a free band if the attack is successful. Besides PUEA, attacks such as covert cyber-deception attacks (CCDA), lion attack, key depletion attack and their countermeasures are described along with cyber-security core concepts in the context of CRNs attacks [30].

The PUE attack's impact is influenced by several variables, including the attacker's position and the CRN's sensing system. Choosing the best spot for the attack will result in numerous secondary users reporting the presence of a primary transmission, which will prompt the CRN to search for a different part of the spectrum. The fake signal, on the other hand, must meet several requirements concerning frequency, code, modulation, etc., to pass as a genuine one if the sensing mechanism used by the CRN searches for specific characteristics of the signal, as is the case with the PN5 sequence, pilot detector, or cyclo-stationary detection. As a result of the attacker's ability to configure its CR device to match the transmission parameters of a primary user or even send a real primary signal that has already been recorded, the PUE attack is more difficult in this situation to implement but is still quite possible [18].

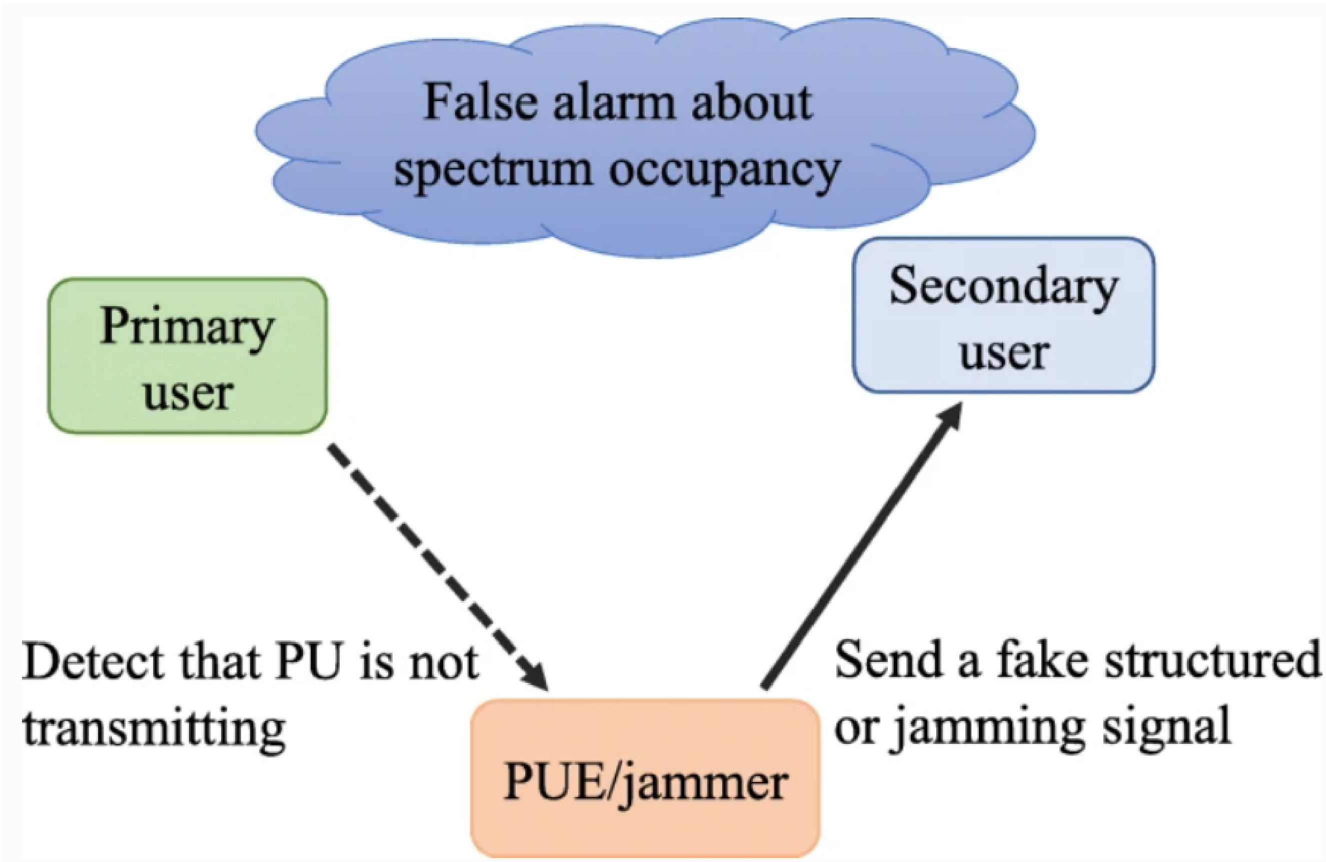

Figure 3: Primary user Emulation (PUEA) and Jamming Scenario.

The identification of PUEA by secondary users is our next goal. For example, we have made the detection time inversely related to the distance between the secondary and malicious primary users. This study uses sparse recovery convergence patterns for PUEA and jamming attack detection. Over a reliable PU channel-dependent dictionary, sparse recovery was made. As a result, the legal node signal exhibits smooth convergence compared to the illegitimate node signal. This attests to the fact that this signal can only be compressed in the domain defined by this sparsifying dictionary. Additionally, the non-compressive properties of a sparse coding jamming signal over a PU channel-dependent vocabulary were used to identify jamming attempts. Based on ML, this detection technique was used [24].
Threats arise due to the wireless nature and unique characteristics of CR. Traditional security threats include eavesdropping, spoofing, and jamming attacks, while new security threats include spectrum sensing data falsification (SSDF) and primary user emulation attacks (PUEA). To assess the effectiveness of CRN in the face of a PUEA attacker, hard decision fusion rules like AND, OR, and K out of N are used. Additionally, a unique reputation-based strategy is suggested, and these fusion rules are employed to determine the final decision to lessen the impact of an SSDF attack. For local sensing, energy detection technology is used [13].

4.7 Spectrum Sensing Data Falsification (SSDF):

In an SSDF attack, a malicious node gives fake sensing data to reduce the effectiveness of a collaborative spectrum sensing strategy. Collaborative approaches involve the interaction of several CRs to enhance sensing performance in fading environments. PUEA, on the other hand, is built on simulating the PU transmission's characteristics to mislead the SUs regarding spectrum occupancy. PUEA inhibits them from utilizing the available spectrum holes and, in some circumstances, may even interfere with the PUs. Various methods have been applied for SSDF mitigation, i.e. from Machine learning. Combating the SSDF attacks is done sifting and evaluation trust management (SETM) algorithm, consisting of two phases, the sifting phase and the evaluation phase in the study as referred here, [9].
In this study referred to at the end of the paragraph, an artificial neural network (ANN) is designed for both PUEA and SSDF attacks and five additional classifiers, including the support vector machine (SVM), random forest (RF), K-nearest neighbours (KNN), logistic regression (LR), and decision tree (DR) for the SSDF attack. Additionally, deterministic and stochastic attack plan scenarios with white Gaussian noise (WGN) are also considered, [39]. This study examines the Probabilistic SSDF attack, a novel attack on the soft decision spectrum sensing paradigm. The attacker in this assault can adaptively alter its parameters. We presumptively receive quantized observations from cognitive users. An honest user reports to FC the quantized level matching its observation. However, an attacker might make false observations and probabilistically translate them onto subsequent quantization levels. [29].
The attacker wants to perform probabilistic mapping to reduce FC's performance as much as feasible. The attacker also takes the cost of the attack into account. To select the best attack parameters, the attacker must solve a convex linear programming problem based on our formulations[2].

5 Discussions

The approach for this stud has been bottom-up research, analysis and simulation-based system design. The study kicked off with the anatomy of the architecture of CR's as an extension of SDR's. After presenting the definition and functionality of Cognitive Radio Networks (CRN's), the study drove towards the architecture of the CRs, the sensing cycle, and the operation of spectrum sensing and acquisition for the secondary users in the empty time-slot of incumbent users.

In the previous Section (Section **??**)

5.1 Accomplishments

The primary objective of the work was simulation bases SS and CRN analysis. The spectrum sensing using co-operative spectrum sensing and Non-co-operative spectrum sensing methods has been studied. The literature is reviewed from the past years' papers on the subject matter, i.e. spectrum sensing using the CRNs and the threshold theory behind its operation and mathematical techniques are covered. The focus, however, has been simulation-based study and real-world scenarios creation-based learning instead of concrete and hard-coded mathematical equations-based derivation-based analysis. After the deep learning penetration in almost every scientific branch, the cognitive cycle could not avoid the deep learning-based system implementation of spectrum sensing systems. Hence, the deep learning techniques mainly used for spectrum sensing, energy detection and opportunistic-based resource allocation using CRNs are covered. Almost all the primary deep learning and machine learning techniques found rich applications in spectrum sensing and cognitive cycle development based on data training methods.

5.2 Key Features of Accomplishments

The spectrum sensing cycle could not escape from the ever-growing threats of weal security in physical, network and application layers of communications. Hence, the cyber-security attacks in the CRN and SS are studied, main attacks are simulated, critical parameters are plotted observed, and conclusions are drawn for well-known attacks in the realm of SS and CRN's such as DoS, SSDF, PUEA ad Node failure. The example topologies are created and results are presented in tabular form.

6 Conclusions

The emphasis has been on experimental results and simulation-based results extraction. Many simulation environments, such as NS-3, MATLAB, NetSim and python, are covered. NS-3 is based on C++ and python. Using NS-3, some simulations have been made, and a network of 10 nodes in data transmission is simulated. The channel is changed to Wimax, and data transmission between the nodes is observed.MATLAB is used for energy detection and spectrum sensing methods simulations. Finally, NetSim is used to create specific scenarios for the cyber security attack simulations in the CRN environments.

Regarding energy detection methods, there are trade-offs, such as MFD requiring complete characteristics of the original signal, while CSD is highly complex in behaviour; machine learning techniques, such as reinforcement learning, are helpful for detection based on the training of the data.

Finally, deep learning techniques benefit DoS, SSDF, and PUE attack prevention. This is because data can train the classifiers or the model from deep learning to detect the attack patterns and send false alarms to automatically prevent the attack, such as a malicious user, PUEA, resource denial or DoS attack.

The simulation-based research can be improved based on the time and resources available for the project. The time constraints and procurement of simulation software proved to be the main barriers to state-of-the-art results compilations. Multiple data sets and data types such as video and audio can be transmitted, however, the main problem with video transmission is that it requires a wide bandwidth communication pipeline, which is rarely the case in Internet-of-things, WSNs, CR-loT or CR-WSN'S. The WSNs are inherently bandwidth and power-limited networks. Due to this constraint, the video data transmission and reception were not realized. The other reason was that NS-3 simulators mainly deal with networking domains, the transmission based on routers and respective nodes, that do not transmit video data except in the case of wired networks for broadband internet access which is mostly fibre-optics based. The basic difference between video data transmission and packet transmission limited the video data transfer and respective delay, and throughout plots.

In this report we presented

7 Annex 2: Acronyms

AODV Ad-hoc On-demand Distance Vector

CRN Cognitive Radio Network

SDR Software Defined Radio

CR Cognitive Radio

ED Energy Detector

CSD Cyclo-Stationary based detection

OSS Opportunistic Spectrum Sharing

DoS Denial-of-Service

MFD Matched Filter Detector

CSS Collaborative/Co-operative Spectrum Sensing

PUEA Primary User Emulation Attacks

SSDF Spectrum Sensing Data Falsification

BLE Bluetooth Low Energy

DSA Dynamic Spectrum Access

RL Reinforcement Learning

TCP/IP Transmission Control Protocol/Internet Protocol

ML Machine learning

ML Long Short-Term Memory

DL Deep learning

RNN Recurrent Neural Network

CNN Convolutional Neural Network

AI Artificial Intelligence

SS Spectrum Sensing

DSA Dynamic Spectrum Access

WSN Wireless Sensor Networks

PU Primary user

SU Secondary user

DNN Deep Neural Network

ANN Artificial Neural Network

Bibliography

[1] *Cognitive Radio Networks*. John Wiley Sons, Ltd, 2009, ch. 6, pp. 145–181. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470742020.ch6

[2] A. Ahmadfard, A. Jamshidi, and A. Keshavarz-Haddad, "Probabilistic spectrum sensing data falsification attack in cognitive radio networks," *Signal Processing*, vol. 137, pp. 1–9, 2017.

[3] S. Akin and M. C. Gursoy, "Effective capacity analysis of cognitive radio channels for quality of service provisioning," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3354–3364, 2010.

[4] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Physical communication*, vol. 4, no. 1, pp. 40–62, 2011.

[5] C. Albon, *Machine learning with python cookbook: Practical solutions from preprocessing to deep learning*. " O'Reilly Media, Inc.", 2018.

[6] Q. I. Ali, "Simulation framework of wireless sensor network (wsn) using matlab/simulink software," *MATLAB-A Fundamental tool for scientific computing and engineering applications*, vol. 2, no. 1, pp. 263–284, 2012.

[7] M. K. Baek and J. Y. Kim, "Effective signal detection using cooperative spectrum sensing in cognitive radio systems," in *2009 11th International Conference on Advanced Communication Technology*, vol. 3. IEEE, 2009, pp. 1746–1750.

[8] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[9] D. Chaitanya and K. M. Chari, "Defense against puea and ssdf attacks in cognitive radio networks," in *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*. IEEE, 2016, pp. 1–5.

[10] N. Chaudhary and R. Mahajan, "Spectrum sensing techniques in cognitive radio networks: challenges and future direction," in *Proceeding of Fifth International Conference on Microelectronics, Computing and Communication Systems: MCCS 2020*. Springer, 2021, pp. 451–458.

[11] Q. Dong, Y. Chen, X. Li, and K. Zeng, "Explore recurrent neural network for pue attack detection in practical crn models," in *2018 IEEE International Smart Cities Conference (ISC2)*. IEEE, 2018, pp. 1–9.

[12] ——, "A survey on simulation tools and testbeds for cognitive radio networks study," *arXiv preprint arXiv:1808.09858*, 2018.

[13] H. M. Furqan, M. A. Aygül, M. Nazzal, and H. Arslan, "Primary user emulation and jamming attack detection in cognitive radio via sparse coding," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, p. 141, 2020.

[14] P. Gawłowicz and A. Zubow, "ns3-gym: Extending openai gym for networking research," *arXiv preprint arXiv:1810.03943*, 2018.

[15] ——, "Ns-3 meets openai gym: The playground for machine learning in networking research," in *Proceedings of the 22nd International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, 2019, pp. 113–120.

[16] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc.", 2022.

[17] G. Ghosh, P. Das, and S. Chatterjee, "Simulation and analysis of cognitive radio system using matlab," *International Journal of Next-Generation Networks*, vol. 6, no. 2, pp. 31–45, 2014.

[18] I. Gupta and O. Sahu, "An overview of primary user emulation attack in cognitive radio networks," in *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, 2018, pp. 27–31.

[19] A. Haldorai, J. Sivaraj, M. Nagabushanam, M. Kingston Roberts *et al.*, "Cognitive wireless networks based spectrum sensing strategies: A comparative analysis," *Applied Computational Intelligence and Soft Computing*, vol. 2022, 2022.

[20] M. Ibnkahla, *Cooperative cognitive radio networks: The complete spectrum cycle*. CRC Press, 2014.

[21] F. K. Jondral, "Software-defined radio—basics and evolution to cognitive radio," *EURASIP journal on wireless communications and networking*, vol. 2005, no. 3, pp. 1–9, 2005.

[22] N. Kassri, A. Ennouaary, S. Bah, and H. Baghdadi, "A review on sdr, spectrum sensing, and cr-based iot in cognitive radio networks," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, 2021.

[23] R. Khana, S. M. Bilalb, M. Othmana *et al.*, "A performance comparison of network simulators for wireless networks," *arXiv e-prints*, pp. arXiv–1307, 2013.

[24] O. León Abarca, "Contributions to the security of cognitive radio networks," 2012.

[25] T. Long and W. Juebo, "Research and analysis on cognitive radio network security," *Wireless Sensor Network*, vol. 2012, 2012.

[26] N. Mansoor, A. Muzahidul Islam, M. Zareei, S. Baharun, T. Wakabayashi, and S. Komaki, "Cognitive radio ad-hoc network architectures: a survey," *Wireless Personal Communications*, vol. 81, pp. 1117–1142, 2015.

[27] J. Mitola and G. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, 1999.

[28] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE personal communications*, vol. 6, no. 4, pp. 13–18, 1999.

[29] N. Parhizgar, A. Jamshidi, and P. Setoodeh, "Defense against spectrum sensing data falsification attack in cognitive radio networks using machine learning," in *2022 30th International Conference on Electrical Engineering (ICEE)*. IEEE, 2022, pp. 974–979.

[30] S. Parvin, F. K. Hussain, O. K. Hussain, S. Han, B. Tian, and E. Chang, "Cognitive radio network security: A survey," *Journal of Network and Computer Applications*, vol. 35, no. 6, pp. 1691–1708, 2012.

[31] W. Prawatmuang, *Cooperative spectrum sensing for cognitive radio*. The University of Manchester (United Kingdom), 2013.

[32] R. Sharma, V. Vashisht, and U. Singh, "Modelling and simulation frameworks for wireless sensor networks: a comparative study," *IET Wireless Sensor Systems*, vol. 10, no. 5, pp. 181–197, 2020.

[33] P. Sivagurunathan, P. Ramakrishnan, and N. Sathishkumar, "Recent paradigms for efficient spectrum sensing in cognitive radio networks: Issues and challenges," in *Journal of Physics: Conference Series*, vol. 1717, no. 1. IOP Publishing, 2021, p. 012057.

[34] S. Solanki, V. Dehalwar, J. Choudhary, M. L. Kolhe, and K. Ogura, "Spectrum sensing in cognitive radio using cnn-rnn and transfer learning," *IEEE Access*, vol. 10, pp. 113 482–113 492, 2022.

[35] B. Soni, D. K. Patel, and M. López-Benítez, "Long short-term memory based spectrum sensing scheme for cognitive radio using primary activity statistics," *IEEE Access*, vol. 8, pp. 97 437–97 451, 2020.

[36] M. Subhedar and G. Birajdar, "Spectrum sensing techniques in cognitive radio networks: A survey," *International Journal of Next-Generation Networks*, vol. 3, no. 2, pp. 37–51, 2011.

[37] B. Suseela and D. Sivakumar, "Non-cooperative spectrum sensing techniques in cognitive radio-a survey," in *2015 IEEE Technological Innovation in ICT for Agriculture and Rural Development (TIAR)*. IEEE, 2015, pp. 127–133.

[38] A. A. Tabassam, M. U. Suleman, S. Kalsait, and S. Khan, "Building cognitive radios in matlab simulink—a step towards future wireless technology," in *2011 Wireless Advanced*. IEEE, 2011, pp. 15–20.

[39] S. Tephillah and J. M. L. Manickam, "An setm algorithm for combating ssdf attack in cognitive radio networks," *Wireless Communications and Mobile Computing*, vol. 2020, pp. 1–9, 2020.

[40] A. Upadhye, P. Saravanan, S. S. Chandra, and S. Gurugopinath, "A survey on machine learning algorithms for applications in cognitive radio networks," in *2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*. IEEE, 2021, pp. 01–06.

[41] M. B. Usman, R. S. Singh, and S. Rajkumar, "Stage spectrum sensing technique for cognitive radio network using energy and entropy detection," *Wireless Power Transfer*, vol. 2022, 2022.

[42] H. Venkataraman and G.-M. Muntean, *Cognitive radio and its application for next generation cellular and wireless networks*. Springer, 2012, vol. 5.

[43] E. Weingartner, H. Vom Lehn, and K. Wehrle, "A performance comparison of recent network simulators," in *2009 IEEE International Conference on Communications*. IEEE, 2009, pp. 1–5.

[44] Y. Wu, J. A. Stankovic, T. He, and S. Lin, "Realistic and efficient multi-channel communications in wireless sensor networks," in *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*. IEEE, 2008, pp. 1193–1201.

[45] J. Xie, J. Fang, C. Liu, and X. Li, "Deep learning-based spectrum sensing in cognitive radio: A cnn-lstm approach," *IEEE Communications Letters*, vol. 24, no. 10, pp. 2196–2200, 2020.

[46] F. R. Yu and H. Tang, *Cognitive radio mobile ad hoc networks*. Springer, 2011, vol. 507.

[47] D. Zhang and X. Zhai, "Svm-based spectrum sensing in cognitive radio," *2011 7th International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1–4, 2011.

[48] Y. Zhang, J. Zheng, and H.-H. Chen, *Cognitive radio networks: architectures, protocols, and standards*. CRC press, 2016.

Artificial Intelligence Assisted Automation of PCB Design Process

An article written for the website, with SEO Optimization as Strategy to increase traffic and creae conversion rate per clicks.

1 Introduction:

Artificial Intelligence (AI) has revolutionized many industries, and Printed Circuit Board (PCB) design is no exception. PCB design, the process of creating the layout and schematic of a circuit board, has traditionally been a time-consuming and complex process requiring significant expertise. However, AI-optimized and automated PCB design has emerged as a solution that streamlines the design process, reduces errors, and increases efficiency. In this article, we will explore the benefits of AI-optimized and automated PCB design and how AI is being used to improve the PCB design process.

1.1 What is Artificial Intelligence?

Artificial intelligence (AI) is a field of computer science that studies how to make computers act like humans. AI can be used for many things, such as creating robots and self-driving cars.
AI uses algorithms to solve problems in ways similar to how humans think about them. For example, if you wanted your robot vacuum cleaner to clean up the dirt on the floor, it would need an algorithm that tells it where all the dirt is on the floor to know what areas need cleaning most urgently.

1.1.1 Machine Learning Model Training Process:

The main steps in artificial intelligence or, more precisely, machine learning/deep learning are gathering and pre-processing the data and splitting the data into training sets, test sets and validation sets. Then a suitable Machine learning model is selected for the desired application, which is trained using the training data set. The machine learning model is trained on data to learn the desired patterns in the data and tune its parameters accordingly. Next time, when it detects the desired patterns, the model can perform the desired operation automatically in terms of finishing the training process. That model can later be used to automate the process and make its own decisions. Hence, the machine learning model training process can be defined below in steps.

1) Data collection and pre-processing
2) Choosing a suitable Machine learning Model i.e. from a given group of available and mostly deployable models.
3) Training and testing the model
4) Model evaluation.

5) Deployment and industrial real world settings

The 2nd step is to have a look at the PCB design process to see which phases can be automated using Machine learning models.

1.2 What is PCB Design?

Before we delve into AI-optimized and automated PCB design, it's important to understand what PCB design is and why it's essential. A Printed Circuit Board (PCB) is an insulating board with conductive pathways etched onto it that connect different electronic components on the board. PCBs are a critical component of almost all electronic devices, including smartphones, computers, and automobiles.

It is essential to look at the design process for a PCB to understand better at which phases it is feasible to implement Machine Learning Methods and hence, the AI-assisted automation of PCB design.

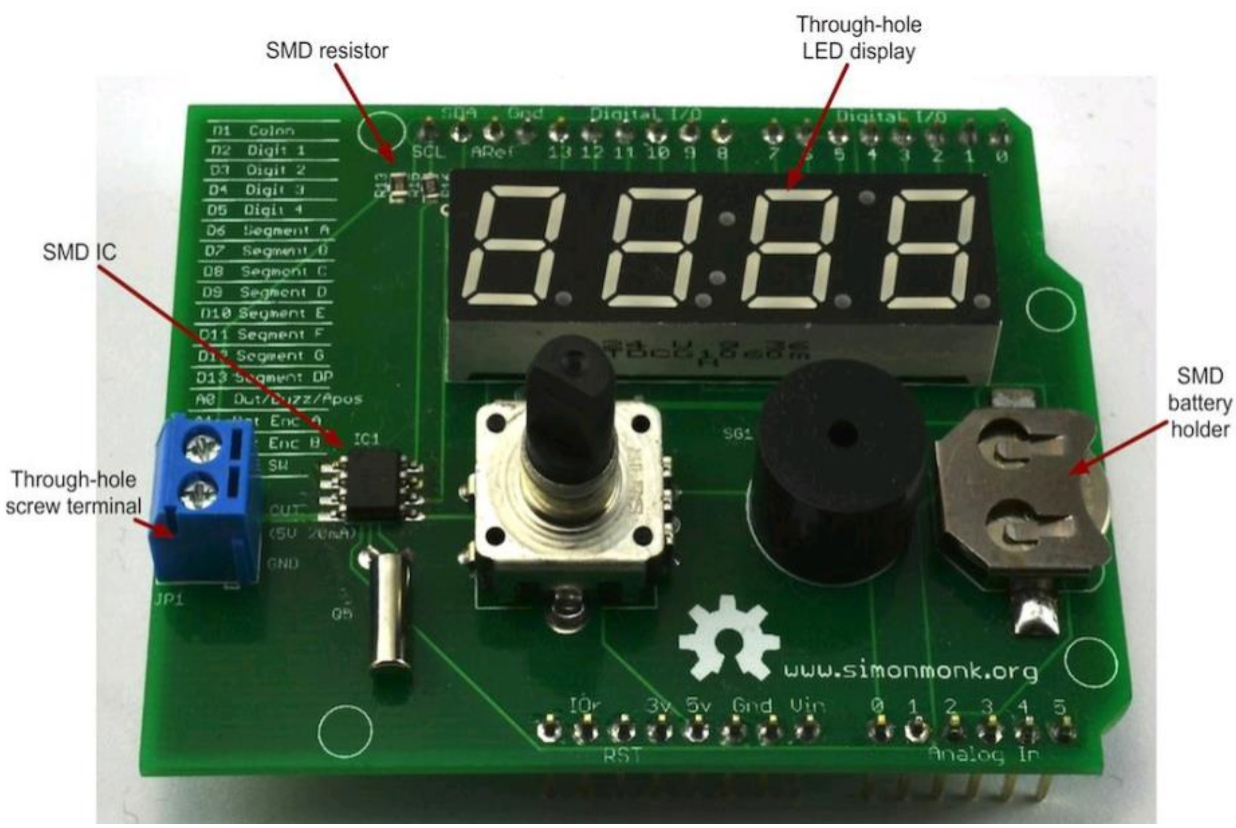

Figure1-1-1: A two layer PcB design

The components are soldered to the PCB at pads.

- Copper tracks that link pads together are known as tracks.

• Vias, which are tiny holes through the circuit board, electrically connect a bottom track and a top track. When laying out a PCB, it is frequently necessary to use a signal to move from one layer to another because tracks on the same layer cannot cross.

• Any lettering appearing on the finished board is called silk screening. To make it simple to see where everything goes when soldering the board together, it is customary to identify components and the contour of where they fit.

• The board is covered on both sides, except for the areas with pads, by a stop mask layer of insulating lacquer.

Surface mount components are soldered to the top of the pad, while through-hole parts are pushed through from the top, soldered underneath, and the excess leads are cut off.

Making a schematic and layout for the PCB, which includes where to position electronic components, how to route electrical connections, and how to maximize signal integrity, is called PCB design. A thorough understanding of electronics, circuit design, and software tools is necessary for the design process [1].

1.3 The Benefits of AI in PCB Design

AI in PCB design can help reduce your design time, improve efficiency, and increase accuracy.

- Reduced Design Times: AI can learn from previous designs and analysis data which means it can use this knowledge to speed up the creation of new PCBs. This allows you to save time by not having to manually create each step in your design process every time you start a new project. The result is that you will be able to get more done in less time!

1.4 Why is AI-Optimized and Automated PCB Design Necessary?

- The traditional PCB design process is manual and time-consuming, requiring significant expertise and experience. Even experienced designers can spend several hours or even days designing a complex PCB. This process can lead to errors, delays, and increased costs.

- AI-optimized and automated PCB design is a solution that can streamline the design process, reduce errors, and increase efficiency. With AI, designers can automate repetitive tasks and leverage machine learning algorithms to optimize the PCB design process. This results in faster design times, fewer errors, and lower costs.

1.4.1 Benefits of AI-Optimized and Automated PCB Design

- AI-optimized and automated PCB design has several benefits, including:
- Reduced Design Time: AI can automate repetitive tasks and suggest optimal component placement and routing, resulting in faster design times.
- Fewer Errors: AI can analyze data and predict potential issues, reducing the likelihood of errors in the final design.
- Improved Performance: AI can optimize signal integrity and select the optimal components, resulting in better performance.
- Lower Costs: Faster design times, fewer errors, and improved performance all result in lower costs.
- More significant Innovation: AI-optimized and automated PCB design frees up designers

1.4.2 AI Automation in PCB Design

There are various ways AI is being used to improve PCB design. Here are some of the most popular methods: AI automation in PCB design can be used to automate the following processes:

- Design Automation. This is the process of using an AI system to automatically generate a circuit design, which is then verified by another set of algorithms before being sent off for fabrication. The main advantage here is that it reduces human error and speeds up production time significantly (by as much as 50%).
- Automated Testing. This refers to using an AI system that can test PCBs automatically without any human intervention required--you just send your files over and let it do its thing! This saves you both time and money because there's no need for manual labor or expensive equipment like oscilloscopes or other testing tools anymore; instead, you only have to wait for results from your computer screen (or phone).
- Automated Production/Manufacturing Processes - One area where this technology has already been implemented successfully involves automating production processes so they run smoother than ever before while also reducing costs significantly over time due to

decreased material wastage rates due to fewer errors caused by human input mistakes during assembly lines where workers manually assemble components into finished products such as computers or smartphones etc.

- Automated Placement and Routing: One of the most time-consuming parts of PCB design is component placement and routing. AI can automate this process by using machine learning algorithms to suggest the optimal placement and routing of components. This results in faster design times and fewer errors. In this phase, the routing and placement by AI have positive results. High-frequency signals can be efficiently routed in multi-layer PCBs using AI-enabled auto-routers. They get knowledge from earlier design examples and construct using human intelligence.
- Signal Integrity Optimization: Signal integrity is crucial in PCB design. AI can be used to optimize signal integrity by predicting and analyzing potential signal integrity issues. This helps designers identify potential problems before they occur, saving time and reducing costs.
- Virtual Prototyping: AI can be used to create virtual prototypes of PCB designs, which allows designers to simulate the circuit's behaviour and identify potential issues before physical prototypes are built. This saves time and reduces costs.
- Component Selection: AI can be used to analyze data from different suppliers and select the optimal components for a given PCB design. This results in better performance and lower costs.
- Design Rule Checking: AI can automate design rule-checking (DRC), ensuring that the design complies with industry standards and best practices. This reduces errors and improves the quality of the final design.

1.5 AI-Powered Design Tools

AI-powered design tools are already available to help you design PCBs faster and more efficiently. The most common use of AI in PCB design is for simulation software, which uses machine learning algorithms to predict how a circuit will behave under different conditions. The software can be used to test the performance of your board before you build it, saving time and money by identifying issues early on in the process.

AI also plays a vital role in other types of design software, such as layout tools, verification tools, and even manufacturing automation systems (MES).

1.6 Integrating AI into PCB Design

Integrating artificial intelligence (AI) into a PCB design process is exciting. It can help you automate tedious tasks and optimize your designs to increase efficiency and performance. Here are some ways that demomstate integrating AI into PCB design processes could improve the quality of your product:

1.7 Utilizing AI for Automation:

With the help of machine learning algorithms, computers can learn how to perform tasks without being explicitly programmed how to do them. This can significantly reduce human error by eliminating human bias from complex decision-making processes such as circuit board layout optimization or automated routing tools.

Leveraging AI for Design Optimization: In addition to automating certain aspects of PCB design workflows, leveraging machine learning capabilities will allow you to access powerful new analysis tools, such as neural networks, to analyse existing data and predict future trends based on historical information.

1.7.1 The value of Experience in PCB design versus Artificial Intelligence, demonstrated by a specific example of selecting the number of layers for a PCB design:

The extraction of certain pieces of information, such as extracting the number of layers in a PCB design, depends on experience. Machine learning or Artificial intelligence should not be of help in this highly complex process. Subsequently, there are many such steps for which the formulation of an exact algorithm may seem impossible due to the incorporation of many factors and intricate calculations and the vast requirement of skillsets and knowledge.

Let's see how PCB design automation is made possible by an example. The PCB design for a particular server or network device comprises the signal, ground (GND), and power supply layers. In the case of the manual design method, the number of signal layers are strategized by a skilled designer historically in PCB designs. This task of estimating the number of layers is attempted to be calculated by applying machine learning methods. This process is termed as artificial intelligence Automation of PCB design, although it is just a sub-process of PCB design procedures that is automated.

Which methods from Machine Learning are known to be useful for specific tasks & different stages in the PCB design process? The Machine learning methods are classified as Supervised learning and unsupervised learning.

In machine learning, the support vector regression from supervised learning is applied to estimate the number of layers. This estimation must be done early in the board design process. Hence the feature vectors in the numerical form are obtained at the early stage of the circuit design process that can be used as the training data for creating the learning model.

The accuracy can be improved for the estimation of the number of layers which is an excellent example for conducting the feasibility study of the machine learning application for the problem of determining the number of layers in PCB design.

The first step in this study is selecting the training data from the design data. The only items chosen for consideration are those close to the estimated target product. The data was split in product type and PCB type.

The second initiative was to review the feature vectors to be input as training data. From a survey of skilled PCB designers, only those items were deduced that were important for the estimation of the number of layers; hence a review of the feature vectors was conducted accordingly.

The third initiative was algorithm optimization for the number of layer estimation learning model creation. For example, the estimation accuracy was improved by changing the kernel function for the support vector regression from linear to non-linear and the parameter supplied to the kernel function [2].

2 Industrial corporations working in AI-assisted PCB design

Zuken is a leader in the implementation of AI in PCB design Technology. As per Zuken, The following are some instances of PCB design activities assisted by AI:

- Optimizing the arrangement of components
- Reduced trace length and improved routing
- Prediction and analysis of signal integrity
- Analysis of dynamic voltage drops
- Checks for DFM (Design for Manufacturability)
- Checking and fixing design rule violations
- 3D modelling and interactive positioning
- Analysis and failure prediction using machine learning.

PCB designers may produce cutting-edge designs with greater functionality while speeding up the design process and improving design quality [3].

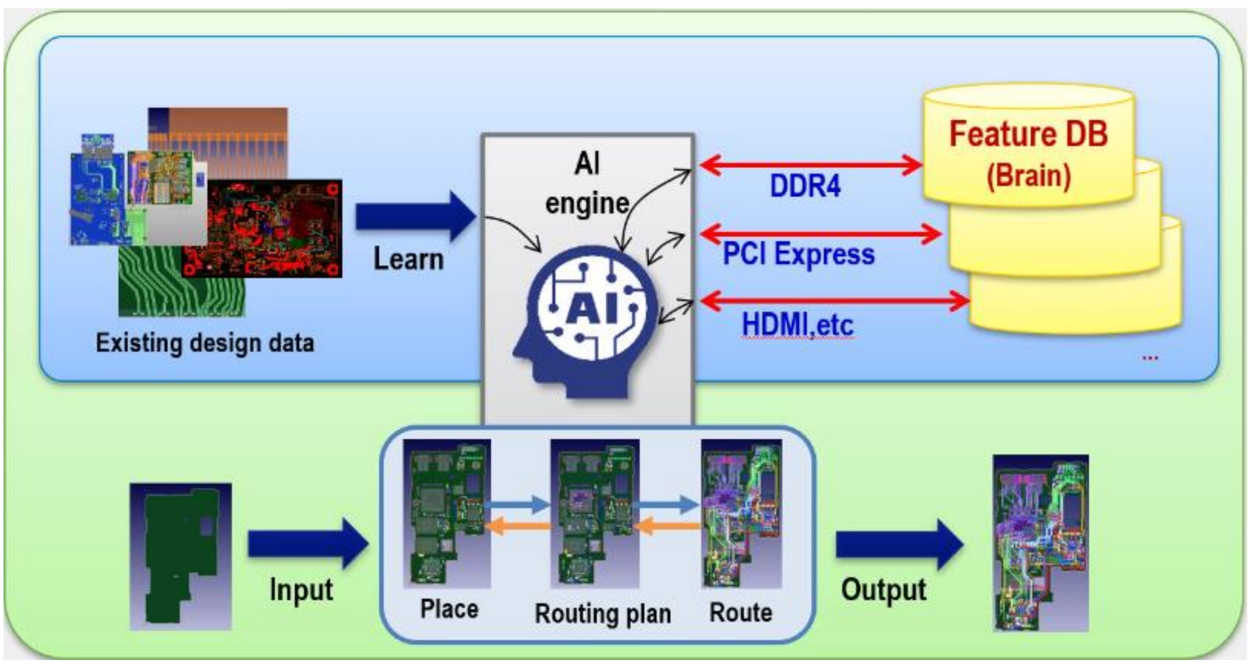

Figure 2-1: A look into the future AI-Based PCB design, Image courtesy by Zuken [3]

However, strides are being made by players such as Zuken, which is AI based PCB Design Company based in JAPAN. Zuken tool for PCB design is called CADSATR, and it provides CADSTAR PCB design editor, Library editor tools (World's largest Libraries in PCB design) and a free PCB design & schematic editor, without any need for license. Some other companies that use AI in

PCB design are JITX, based in USA, Continuity, NOW Celus, based in Germany, Circuit Tree, founded in INDIA, Circuit Mind, Based in UK, Instadeep, based in London and Gumstix Geppetto, based in USA. An introduction to these five companies and their respective progress in AI-based PCB design milestones is described in the article titled as "5 Printed Circuit Board Design Tools That Use AI" [4], [5]. To build a comprehensive library of Cubos, CELUS has relationships with companies like Avnet, Kyocera AVX, Molex, and Würth. The Cubos' flagship Supernova design tool enables designers to create new products quickly [6].

Instadeep PCB design platform is called deep PCBTM beta 8.0. A fully automated tool called DeepPCBTM is now in beta testing. Reinforcement Learning (RL), an AI technology particularly suited for decision-making challenges, such as board games like Chess or Go, logistics, mobility, or PCB routing, makes automation possible. At InstaDeep, RL systems are developed in the real world and in collaboration with Nvidia and Intel (we are a member of Intel's AI Builder Program). As per programs manifest, there are no people in the process to reduce human errors and increase efficiency, which helps accelerate PCB development cycles. Users now have access to intermediary routing solutions, demonstrating that DeepPCBTM is entirely automated. The user may observe how, as the AI system gains experience solving the user's board, the quality of its routing gradually improves.

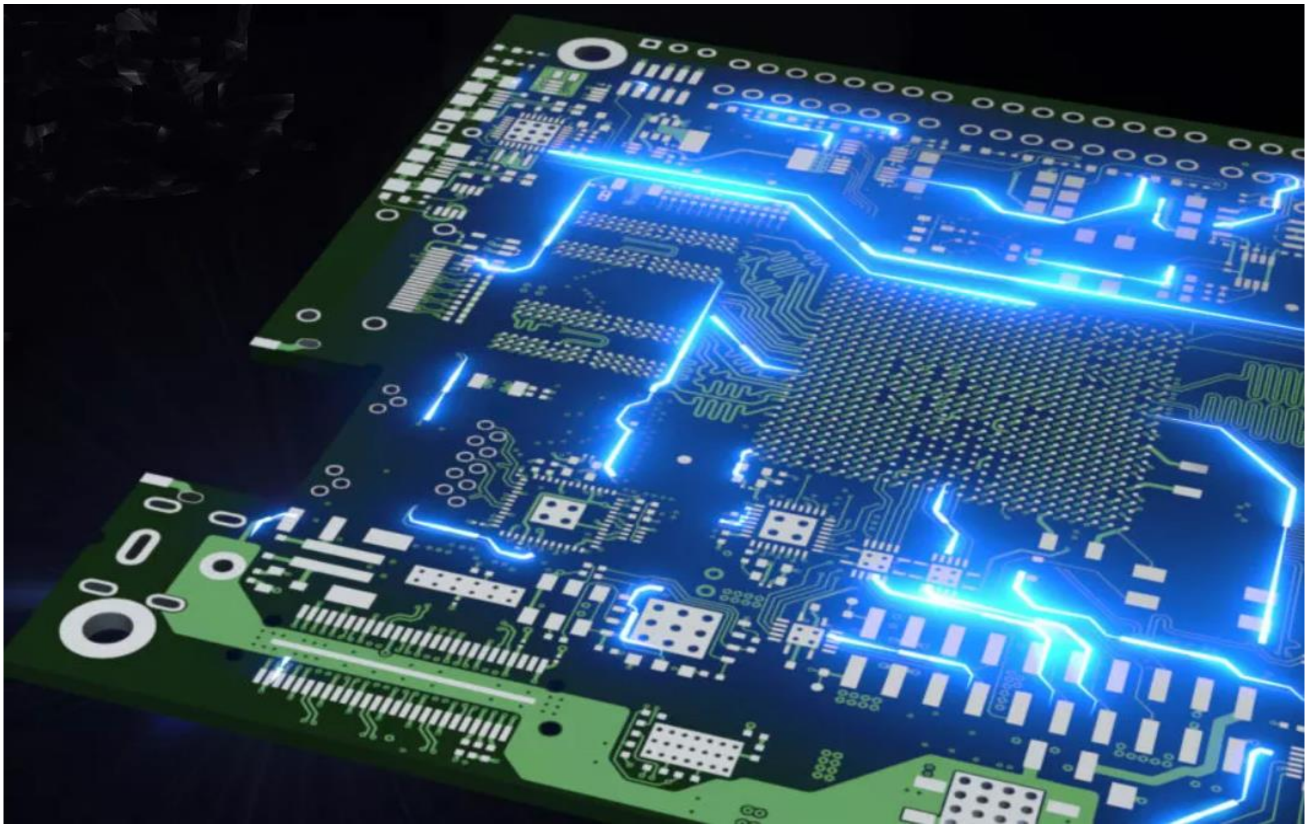

Figure1-2: Image courtesy Instadeep

Additionally, Cadence asserts that the Allegro X platform's built-in AI technology can 10x reduce production turnaround time. Their Software, known as Allegro X AI, which is integrated into the Allegro X platform for system-level design by the Santa Clara, California-based business, can autonomously develop physical layouts of small to medium PCBs. As a result, compared to a manually created deisgn, it aids in focusing on layouts of equal to superior quality.

2.1 The Future of AI in PCB Design

There are many other uses for AI in PCB design. These include automated testing, production and design automation.

Automated Testing: One of the most exciting areas for AI is how it can be used to test PCBs. Traditionally, this has been done manually by designers going through each board individually, checking for faults and errors before sending them off for manufacture or assembly. This process can take days or even weeks if there are a lot of boards to check over (for example, if you need to test several hundred units). With an automated system, however, this could be reduced to minutes or even seconds! The benefit here is obvious; you save time and money, too, since there's no longer any need for someone else's labour costs (which may add up quickly depending on how many prototypes need checking).

2.2 Conclusion

AI is already a part of our daily lives, from the music society listens to, the news read daily by individuals, and how shopping is done online. AI has also made its mark on PCB design. It's been around for years--but only recently has it become more accessible and affordable for designers. AI automation in PCB design is just one example of how AI can help save time while increasing productivity and quality control standards. Another way to integrate artificial intelligence into PCB design workflow is through AI-powered design tools such as Cadence Allegro or Mentor Calibre (formerly known as Mentor Graphics PADS). These programs use machine learning algorithms that allow them to learn from previous designs to predict potential issues before they happen--saving designers even more time!

The future looks bright for this technology: experts predict that by 2022 there will be an estimated 1 million jobs created by artificial intelligence alone--and with those numbers, there's no doubt that more companies will begin integrating this technology into their processes soon enough!

However, the integration of AI in PCB is still in the infancy stage. Progress is being made is developing databases for the Routing phase for two-layer to multi-layer (10 or more) PCB designs and the complex design rules check (DRC) phase. However, these are complex procedures, and there is a lack of data to train for these complex skill-based tasks.

3 References

[1] Monk, S., & Amos, D. (2017). *Make your own PCBs with EAGLE: from schematic designs to finished boards*. McGraw-Hill Education.

[2] Nozaki, N., Konno, E., Sato, M., Sakairi, M., Shibuya, T., Kanazawa, Y., & Georgescu, S. (2017). Application of artificial intelligence technology in product design. *Fujitsu scientific & technical journal*, *53*(4), 43-51.

[3] Zuken Blog, A chat with ChatGPT, Is using AI for printed circuit design a realistic proposition? https://www.zuken.com/en/blog/ai-pcb-design-a-chat-with-chatgpt/

[4] 5 Printed Circuit Board Design Tools That Use AI, Online, and Available at: https://www.rtinsights.com/5-printed-circuit-board-design-tools-that-use-ai/

[5] Automated PCB Design Using Artificial Intelligence (AI), Online, and Available at:

https://pallavaggarwal.in/automated-pcb-design-using-ai/.

[6] Dale Wilson, CELUS Brings Modern Concepts and AI to Circuit and PCB Design—Exclusive, 12, 08, 22, Online, Available at: https://www.allaboutcircuits.com/news/celus-brings-modern-concepts-and-ai-to-circuit-and-pcb-design/Accessed [2nd May 2023]

WSNs: Multi-Hop Transmissions using Evolutionary Algorithm

ABSTRACT

Data can be gathered, analyzed, and collected via Wireless Sensor Networks (WSN) set up in a particular field. Later, through some routing algorithms and techniques, this information can be sent from the nodes to the base station using various communication techniques and application layer protocols. Constrained Application Protocol (CoAP), which is based on User Datagram Protocol (UDP), Message Queuing Telemetry Protocol (MQTT), which employs the Transmission Control Protocol (TCP) transmission mechanism, Extensible Messaging and Presence Protocol (XMPP), which also employs TCP, and Representational State Transfer (REST) based on Hyper Text Transfer Protocol (HTTP) are some examples of application layer protocols. These protocols can be applied to transport data between the sensors and base station, cloud, gateway or any other highly-processing capable device. The overhead or payload for each protocol varies; hence, energy consumption can be reduced in data transmission through routing protocol modifications and improvements. Besides, specific protocols, such as TCP-based protocols, provide better security in the physical layer. There are also techniques for clustering the nodes and transmitting data from nodes to Cluster Heads (CHs) and then to the base station (BS).

However, a bottleneck in Internet-of-thing (IoT) based WSN deployment is their limited bandwidth, battery power, and computing power capability. The primary reason for low energy consumption and efficiency requirement is that WSNs are deployed in remote, inaccessible locations and left alone and hence are difficult to recharge; 2nd wireless or IoT sensors are inherently power limited, as designed by manufacturers. This reduces the network lifetime and makes it a critical issue along with energy management. Now, a sensor captures data, processes it, and must send it to the nearest location, such as a base station, edge processor, fog network, or cloud. It has to ship because of the 2nd major sensor limitation, which is limited onboard processing capability. It has to rely on data transmission to a more intelligent and capable device operating nearby, primarily a base station in wireless sensors and remote deployments. This is because the cellular infrastructure is the only one that is widespread geographically and provides enough coverage. Hence, the transmission of the data by a routing protocol is the most power-consuming part, which is why intelligent algorithms are needed and optimization strategies to be developed to conserve the energy of sensors.

Many algorithms have been researched, such as LEACH (low power adaptive clustering Hierarchy) and LEACH variants such as LEACH-C, Modified LEACH (MODLEACH), Distributed Energy Efficient Clustering Protocol (DEEC), Power-Efficient Gathering In Sensor Information Systems (PEGASIS), and Threshold sensitive Energy Efficient Sensor Network (TEEN).

In this study, we review the existing literature and analyze several routing algorithms to see how CH selection varies between them. Following an overview of methods influenced by genetics and nature-inspired algorithms, we present game theory routing protocol implementations. We examine and contrast the outcomes of cutting-edge routing algorithms and suggest an algorithm that enhances the earlier research efforts conducted.

Table of Contents

LIST OF TERMS AND ACRONYMS

Acronyms	Terms
MPC	Multi-hop clustering
WSN	Wireless Sensor Networks
IoT	Internet-Of-Things
MAC	Medium Access Control
OSI	Open Systems Interconnection
LEACH	Low Power Adaptive Clustering Hierarchy
IGABACA	Improved Genetic Algorithm and Binary Ant Colony Algorithm
DTN	Delay Tolerant Network
CHs, CH	CHs, CH
LO	Lion Optimization
PEGASIS	Power-Efficient Gathering In Sensor Information Systems
TEEN	Threshold sensitive Energy Efficient Sensor Network
TDMA	Time division Multiple Access
DEEC	Distributed Energy Efficient Clustering Protocol
FND	First Node Dead
HND	Half Node Dead
LND	Last Node Dead

ETASA	Energy And Traffic Aware Sleep-Awake
TEAR	Traffic-Energy Aware Routing
SEEDS	Scalable Energy Efficient Deployment Scheme
REST	Representational State Transfer
XMPP	Extensible Messaging and Presence Protocol
MQTT	Message Queuing Telemetry Transport
CoAP	Constrained Application Protocol
AMQP	Advance Messaging Queuing Protocol

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

1. INTRODUCTION

Chapter 1 formulates the problem statement and covers the introduction to WSNs, the topologies in WSN design, applications of WSNs, and the objective of the project, followed by the scope and significance of the study; finally, it describes the project report outline.

1.1 Introduction to WSNs

Internet-of-things is a concept of connecting every aspect of modern life to the internet. Using sensors, the data consisting of various physical and environmental parameters data is collected and uploaded to a cloud or base station for processing.

WSNs are infrastructure-less networks typically deployed by spreading many sensors in a field where the physical and environmental parameters, such as temperature, humidity, and pressure, need to be measured. Typical applications for WSNs include Internet-of-things, security, and surveillance for threat detection, environmental parameter collections such as temperature, pressure, humidity, patient monitoring, agriculture, and measuring the background noise level(Geeksforgeeks, 2021).

According to some surveys by Intel and Cisco, Internet-of-things devices will be deployed in billions of numbers; to be exact, Intel numbered 200 billion sensor devices while Cisco numbered 50 billion. This shows the proliferation of the WSNs or IoT in almost all segments of life, i.e., commercial, industrial, and medical.

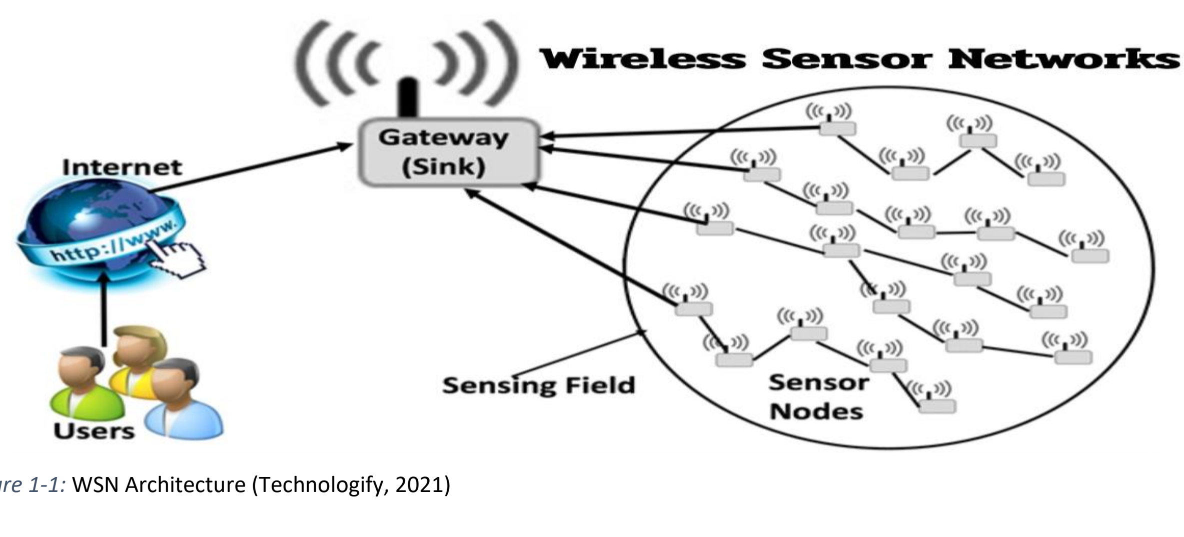

Figure 1-1: WSN Architecture (Technologify, 2021)

1.1.1 Applications of WSNs

If two main application areas are considered, i.e., military and civil, the application methodology can be classified into two areas: data collection and surveillance. However, widespread application areas are as given below.

- Military Applications

Militaries can use WSNs to obtain information from a battlefield for surveillance and reconnaissance and attach sensors to soldiers for their vital signs collection so that physician can handle health emergencies well in advance.

- Zone monitoring

The sensors can be deployed in a particular zone; when physical or environmental parameters such as heat, pressure, and temperature exceed a threshold, appropriate actions can be triggered.

- Transport

Sensors can be attached to drivers, alerting if a driver goes into sleep mode or to the cars giving data about traffic congestion and other traffic issues. Modern autonomous vehicles have many sensors, such as cameras, radars, lidars, position sensors, gyroscopes, etc.

- Industrial Monitoring

Industrial WSNs can be used to give data about machines and equipment to provide better maintenance, repair, and troubleshooting activities, saving costs and downtime. Such concepts are predictive maintenance, corrective maintenance, and preventive maintenance (Taylor, 2022).

- Health and well-being monitoring

With sensor technologies such as wearables and bio-medical parameters, it is possible to get the body's physical parameters (vital signs) periodically for continuous health monitoring. This data is transmitted to doctors in real-time and rapid support is ensured in case of emergency using emergency communication methods. The parameters to be monitored are physical and cognitive, such as ECG, EEG, blood pressure, temperature, pulse rate, respirations rate, saturation rates, and pulse oximetry.

❖ Environmental sensing

This application includes the detection of glaciers, and volcanoes, monitoring of air pollution, fire in forest detection, and landslide detection (Technologify, 2021).

1.1.2 Design Issues in WSNs

Due to heterogeneous deployment, there are many hurdles and challenges in implementing WSNs.

i. Because sensors are manufactured to be small in size and have small batteries, they are limited in battery power, bandwidth, and data rate.
ii. The terrain to be monitored is widespread and heterogenous; therefore, sensors are randomly thrown inside the field in large numbers, making each sensor coverage area overlap.
iii. The terrains are rough, and unexpected earthquakes and explosives can happen. Hence, topology has to be reconfigured frequently.
iv. Wireless links are random, and the signals experience fading, shadowing, scattering, and interference from other communication links; hence dropped packets and link disruptions happen frequently.
v. The sensors are placed in a hazardous environment; therefore, they can fail, or the battery power may exhaust over time with no backup for the sensor, and the sensor node may die.

1.1.3 Topologies for WSNs

Multiple topologies can be designed when deploying WSNs in a field, battlefield, or observation space.

❖ Star Network

It is a centralized architecture. Star topology interconnection for WSNs considers a central node, a base station for WSNs, a hub or switch, and a router in networking terms. Each node will require a dedicated channel to communicate with the central hub. Hence, there will be N channels in the case of N nodes. In case the central node or base station collapses, the system goes down.

❖ Mesh Network

Communication between nodes takes place directly. In a mesh, if n nodes participate, the total number of channels will be n(n-1)/2. This network can self-configure or self-organize itself Kiran Kumar Panigrahi, (2022).

❖ Hybrid

Hybrid is an energy-efficient topology, and Zigbee utilizes it. Using multi-hop means that the power consumption of nodes with multi-hop is higher than those with one-hop. The advantage is that low-power nodes can go to sleep to conserve energy while not communicating.

1.2 Problem statement

The concept of WSNs is a crucial step towards digitization; however, serious problems are associated with the successful deployment of WSNs. The sensors are manufactured by keeping in mind the tiny size of the sensors, reduced form factor, and less cost. The sensors' battery life (limited energy efficiency), processing power, Quality-of-Service (in data transmission), lack of privacy and security mechanisms, network throughput, automatic node failure recovery, cross-layer optimization, and scalability issues in large-scale deployments are all limited as a result. Out of these issues, energy efficiency, processing power, and security are controllable through intelligent algorithmic implementations. Multi-hop implies that a field's numerous geographically dispersed nodes share a single wireless medium; as a result, effective Medium Access Control (MAC) design and routing protocols are required to take advantage of spectrum reuse fully. Cross-layer integration is necessary due to the wireless channel complexity and time-varying randomness, which necessitates that the higher-level Open System Interconnection (OSI) layer take the physical layer's features and properties into account. MAC design that is both effective and efficient, choosing the appropriate level of inter-layer communication to create reliable, usable, and more efficient systems, and calculating the fundamental limitations of multi-hop networks under realistic conditions are the challenges.

1.3 Review of existing simulation environments for WSNs

The simulation software for WSNs is varied in use in research labs and companies. MATLAB Simulink can be an excellent choice. MATLAB scripts can simulate the nodes in a WSN and plot their relevant energy consumption, network lifetime, nodes coverage, and clustering-based nodes (CHs and Sink nodes). Some other simulators include GloMoSim/QualNet, OMNeT++, TOSSIM, OPNET Modeler Wireless Suite, MiXiM, Castalia, INET framework, NesCT, Avrora, NS-2, J-SIM, ATEMY, Emstar, SENS, SENSE, and SHAWN., Ali, (2012).

Two scenarios are considered, with 50 nodes in one design and 100 in the other. The simulation is performed with the same parameters on different simulators such as OMNet++, NS-2, NS-3, and MATLAB/Simulink, and the results were surprisingly faster for MATLAB/Simulink. MATLAB exhibited a remarkable decrease of 94.11%, 93.10%, and 92% in simulation runtime when competing with other WSN simulators such as NS-2, NS-3, and OMNet++, Sharma et al., (2020). The parameters are given in Table 1.1, and the results are displayed in Figure 1.2.

Table 1-1: Simulator's performance comparison for WSNs. Sharma et al., (2020)

Simulator	Scenario 1, 50 nodes (ms)	Scenario 1, 100 nodes
NS-2	3.4 ms	3.6 ms
Ns-3	2.9 ms	2.6 ms
OMNet++	2.5 ms	2.7 ms
MATLAB	0.24 ms	0.7 ms

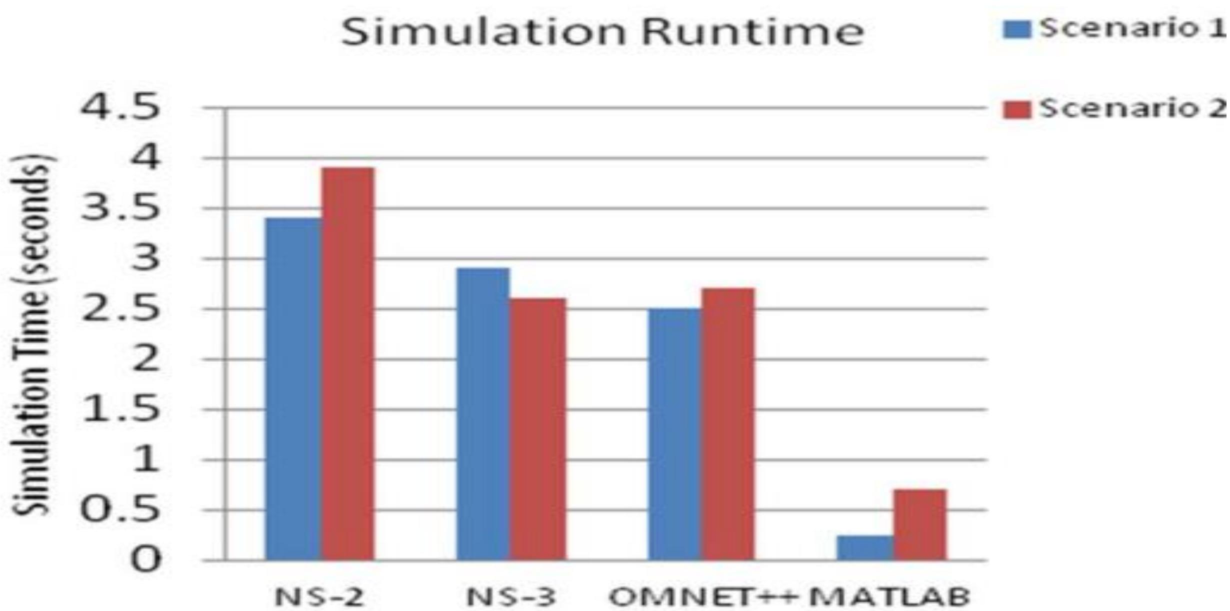

Figure 1-2: Simulation runtime performance comparison of different simulators

Flooding, Directed Diffusion, and LEACH are three protocols that we will be comparing while simulating their data transmission methods on the OMNeT++ platform and offering suggestions for protocol enhancement. The results show that LEACH has the best energy consumption performance when compared to direct diffusion and flooding technique. When the simulation time is identical for all, the LEACH protocol is more trustworthy than the Flooding protocol and the Directed Diffusion protocol. The Flooding protocol has the highest percentage of packages lost, followed by the Directed Diffusion protocol and the LEACH protocol, in that order.

The time needed for events to flow from the source to the sink node in a simulation is known as the WSN time delay. WSNs employ a variety of jumps for data transfer; routing leaps directly impact transmission time delay. Information implosion, overlapping, and resource blindness are problems with the flooding process that result in a noticeable delay. The production of a vast amount of data and control packages, as well as the packages' loss and resending owing to energy constraints in the route establishment stage, may cause the Directed Diffusion protocol's time delay to increase, even though it is less than the Flooding protocol's. The Directed Diffusion protocol has a time delay in the middle as per the results. The Leach protocol speeds up route creation by integrating data into the route-establishing and data transfer processes. Of the three protocols, the Leach protocol has the most minor time delay, Xue & Ren, (2012).

1.4 Objectives

This project aims to work on an algorithm for the energy consumption reduction of the WSNs in the networking layer; hence we focus on the routing algorithms used for the transportation of the data from sensor nodes to the base station, edge computer, cloud, or Fog computing.

The following breakdown of steps can be followed to achieve the objectives of this project:

a) To present a systematic literature review analysis of the recent state-of-the-art paper and works on routing algorithms for data transmission (from the sensor to the base station) in energy efficient manner.

b) To model and optimize a unique routing algorithm and variant of the recent state-of-the-art works from researchers in the past 5-7 years.

c) To implement a novel algorithm that improves the energy consumption, network lifetime, coverage, and related parameter for the WSNs.

d) Simulate the proposed algorithms and compare the results of conventional methods and their improvements for efficient routing protocol methods, i.e., LEACH and MODLEACH, DEEC and its variants. The purpose is to demonstrate improved energy efficiency, network lifespan, dead and alive node and/or delay time between nodes and base stations.

1.5 Background

Routing in multi-hop WSNs (WSNs) is a critical task that involves finding an optimal path for data transmission from a source node to a destination node. Due to the constraints of WSNs, such as limited energy and resources, traditional routing protocols may not be suitable, He, (2009). Evolutionary algorithms (EAs) have been proposed as a solution for routing in WSNs for enhanced power efficiency and improved network lifetime.

One popular EA-based routing protocol for WSNs is LEACH which utilizes a clustering approach to reduce energy consumption, X. Liu et al., (2008). LEACH variants, such as LEACH-C and LEACH-Centralized, have been proposed to improve the original LEACH protocol by addressing such problems as CH selection and data aggregation.

Another EA-based routing protocol for WSNs is TEEN, which uses a threshold-sensitive approach to improve energy efficiency. Power-Efficient Gathering In Sensor Information Systems (PEGASIS) and Stable Election Protocol (SEP) are TEEN variants proposed to strengthen the original TEEN protocol by addressing issues such as chain formation and CH selection, Khan et al., (2015).

Overall, EA-based routing protocols for WSNs such as LEACH, LEACH variants, TEEN, PEGASIS, and SEP, effectively improve energy efficiency and extend the lifetime of WSNs, Nam, (2020). However, further research is needed to address these protocols' scalability and security issues and critical parameters such as energy consumption, network lifetime, and throughput capacity.

1.6 Significance

The significance of a study on routing in multi-hop WSNs using evolutionary algorithms and LEACH variants is that it combines the advantages of two approaches to improving routing performance in multi-hop WSNs. LEACH is a well-known protocol for WSNs that aims to reduce energy consumption by forming clusters of sensor nodes and rotating the CHs. Evolutionary algorithms, such as genetic algorithms and particle swarm optimization, can be used to optimize the parameters of the LEACH protocol to improve its performance Huruială et al., (2010).

This kind of research is essential because it deals with the crucial problem of energy efficiency in multi-hop WSNs, which is a challenge for these kinds of networks. The LEACH protocol's parameters are optimised using evolutionary algorithms in the study to increase the network's scalability, energy efficiency, and overall performance. The results of this study may also lead to new applications for evolutionary algorithms in other kinds of networks and offer guidance for developing WSNs with longer-lasting networks and energy-efficient routing protocols.

1.7 Scope

MATLAB was chosen as the simulation environment because of its speed when compared to other simulators like OMNET++, NS-2, NS-3, and so on.

While researching, an attempt has been made to study previous research papers that mainly considered three models for WSN routing algorithms: the radio channel model, the energy consumption model and the radio interference model. In the energy consumption and radio channel model, assumptions are made by considering a single transmitter and receiver and a channel in between; later, this model was extended to multi-hop multiple WSN nodes. Channel is modelled as well using fist order radio channel modelling theory. Again, a comparison is made between a measurement-based approach and a theoretical-based approach to channel modelling. The channel effects on energy consumption with the distance are studied through simulated results. Another mode, termed aggregation model for CH selection, is also planned for the study as data aggregation at the CH nodes results in higher consumption at that node and a shorter lifetime of CH nodes.

Although network lifetime and energy consumption are the two most crucial factors, It is also necessary to consider the channel capacity, signal-to-noise ratio (SNR), latency, energy efficiency, throughput and packets sent to CH/BS. The energy consumption of multi-hop and single-hop is also compared in various settings.

First, LEACH algorithms are studied and compared to conventional transmission algorithms such as direct transmission. Simulations are made to prove the effectiveness of LEACH compared to direct transmission. Then LEACH variants such as MODLEACH and other clustering protocols such as DEEC, TEEN and other improvements are studied systematically, and results are compared in MATLAB. In the last step, an attempt will be made to experimentally compare the outcomes of genetic algorithms and other nature-inspired algorithms. A unique approach will be suggested if successful in enhancing the results of earlier studies covered in the literature review.

1.8 Project outline

There are four chapters in this report. The introduction of this study, including its context, issue description, aims, scope, and significance, is covered in Chapter 1.

A current literature overview on routing protocols for WSNs is presented in Chapter 2. This chapter addresses several techniques, including clustering-based routing protocols, LEACH and its variants, DEEC variants, nature-inspired algorithms, and game theory-based routing algorithms. There are also reviews of last five years in past and recent studies on the topic. This literature review gives a clearer sense of the strategies academics and researchers have used to approach the issue of reducing energy consumption and network lifetime and what approaches will be taken moving forward to get better results.

The models have been developed for a radio channel, power consumption and energy dissipation modelling in previous studies on this topic; hence, the suitable configuration of each model is chosen for the problem at hand. Other recommended issues related to this topic are described in Chapter 3. This chapter provides specifics and descriptions of the mathematical analysis and parameters.

The outcomes of this study and results are then presented in Chapter 4. In this chapter, the effectiveness of several routing protocols is compared. Finally, a discussion has been made to discuss the progress made so far. Future plans are also discussed. Results and conclusions are described, and recommendations for future directions for this work are described.

Chapter 5 concludes by discussing the goals set for this research project and its contribution to our understanding of cutting-edge routing algorithms used in WSNs. This chapter comes to a close with a discussion of our work's shortcomings and potential directions

CHAPTER 2

2. LITERATURE REVIEW

2.1 Overview

This chapter will discuss the past studies that have been carried out regarding energy-efficient routing algorithms in multi-hop WSNs. The conventional routing protocols (direct transmission, static outing) and their better-performing counterparts, such as LEACH, are discussed. LEACH variants and some more routing algorithms are discussed and compared. The finding of this chapter will be the reference and guidance for this project, Kulkarni et al., (2013).

2.2 History and Background of Routing Protocols in WSNs

LEACH was the initial algorithm for clustering in WSNs proposed by Heinzelman et al., (2000). The idea is based on clustering sensors based on their randomized rotation and forming CH to distribute the load among sensor networks evenly. Localized coordination between sensors and integrated data fusion into the routing protocol reduces the information or overhead that needs to be transmitted to the base station (BS). The even distribution of energy decreases energy consumption. Compared to standard direct transmission, minimal transmission energy, multi-hop routing, and static clustering, it increases energy efficiency and network lifetime.

One interesting heterogeneous method is DEEC for Heterogeneous WSNs, as L. Qing *et al.*, (2006) proposed. The LEACH performance degradation in a heterogenous configuration as compared to a homogenous one is discussed as well in this study. DEEC is a two-level heterogeneous protocol in which nodes are selected to be CHs depending on the likelihood or probability of remaining energy to average energy ratio. Enhanced Distributed Energy Efficient Clustering (E-DEEC) is a three-level approach that assigns another node, termed a super node, and increases the heterogeneity of the network, thus increasing the network's stability and lifetime, aini & Sharma, (2010). Direct communication between nodes and BS is the natural approach to implementing WSNs. However, it has two flaws, the farthest nodes consume more energy and have less signal strength compared to nodes close to BS. LEACH distributes the same energy to each node homogeneously. A better approach will be to heterogeneously distribute more energy to advanced and farthest from BS nodes, termed CH (CHs). The paper Redjimi et al., (2022) compared the performance of DEEC, and E-DEEC protocols and demonstrated E-DEEC outperforms DEEC in the lifespan of the network and packets sent to the BS.

The LECAH protocol has its limitations, though. High packet loss in a congested network is one restriction of the LEACH routing network. This limitation was addressed by, Rahmadhani et al., (2018), and they suggested a solution based on a delay tolerant network (DTN, LEACH-WSNoverDTN), which does not considerably reduce the energy consumption compared to LEACH-WSN but, in network lifetime, showed an early death time of nodes. Hence, LEACH-WSNoverDTN proved to be more efficient as compared to LEACH-WSN in a dynamic network.

The multi-hop clustering (MPC) is another algorithm that uses techniques such as power control at each sensor node, distance calculation from the node to the point to transmit, and directional antennas. Each node has a unique ID and compression/data aggregation capability, and the operation is divided into three rounds, Phase 1: CH selection, phase 2: cluster formulation, and the steady-state phase. CHs' energy usage is reduced using the MPC approach and its practical implementation, Yi et al., (2010).

Each round of LEACH has two phases, the set-up phase and the steady state (stable phase). During the set-up period, clusters are formed; however, nodes do not communicate with the sink node in a steady state.; however, nodes send data to CHs before sending it to the sink node to complete fusion processing. The constant time interval is more significant than the set-up state time duration, ensures reduced power consumption.

LEACH, however, has the following shortcomings.

1. CH selection is a purely random process; hence the node with low energy can be selected that can die sooner.
2. The CH number and location are evenly allocated; this can hurt the lifespan of the network and effect network load-balancing.
3. Since all CHs transmit and communicate to the sink node, if node energies are evenly allocated, the node farthest from the sink node will die away sooner than the nearest nodes.

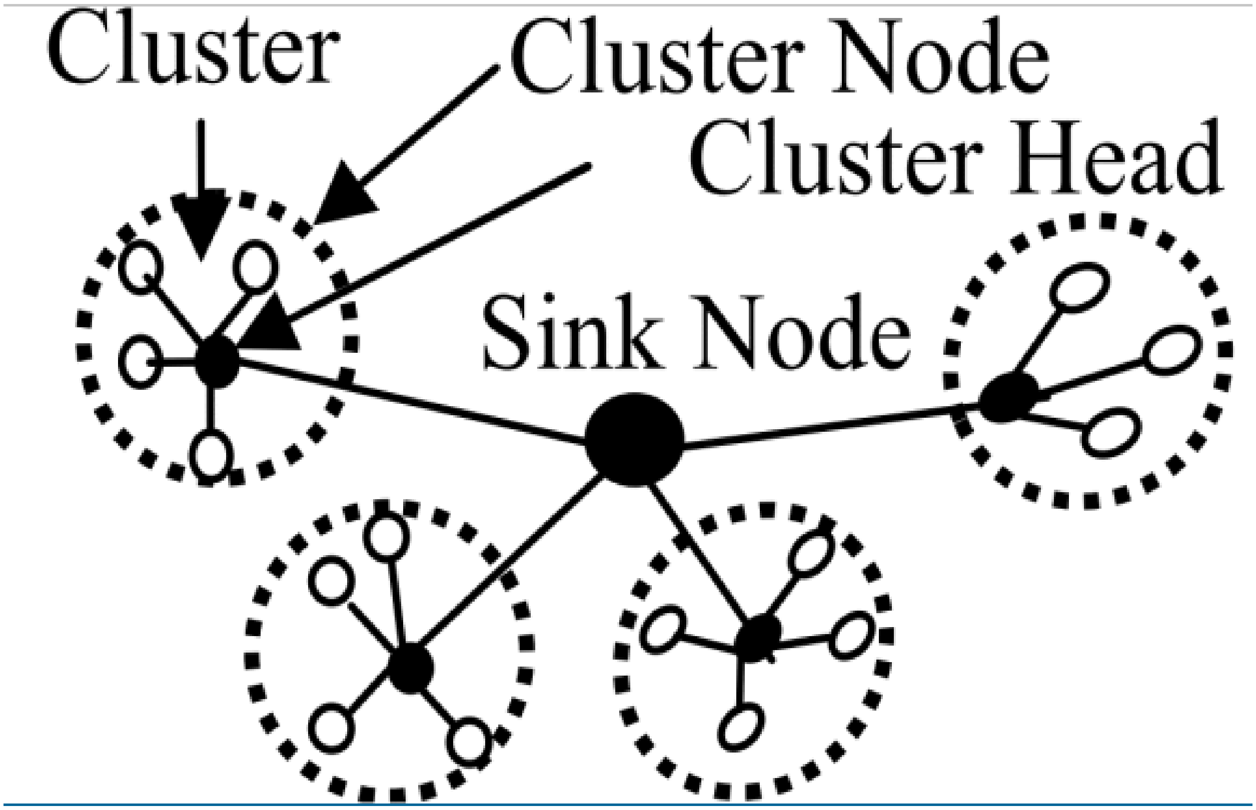

Figure 2-1: LEACH Architecture, Nodes, CH, and sink node (Source: Wang *et al.*, (2009))

Solutions to these problems are addressed by Wang et al., (2009) in the form of a new protocol, the Multi-Hop Routing Clustering Algorithm (MH-LEACH). It reduced energy usage and improved network lifespan, as illustrated by the results in the paper.

LEACH-SM is another protocol that uses the concept of spare nodes in WSN to enhance network lifetime. These nodes are in latent or sleep mode and are only used when some other node exhausts its energy. This configuration adds a spare selection phase to normal LEACH. Hence, how long a spare must be in sleep (passive) mode and which spares are to be used to replace regular nodes are questions addressed in this study., Bakr & Lilien, (2011).

Researching LEACH variants, a very nice comparison of LEACH variants (LEACH, LEACH-C, LEACH-1R, and Hybrid Energy Efficient Distributed Clustering protocol (HEED) is presented by Omari & Laroui, (2015). The authors compared the performance of these clustering-based algorithms with the metrics such as live node numbers, data transmission to sink/BS, residual energy, and the maximum number of rounds. LEACH, LECAH-C, and LEACH-1r are compared.

HEED variants such as HEED-MIN, HEED-MAX, and HEED-AMRP are also compared in performance.

Another variation of LEACH is LEACH-M, based on the distributed address assignment mechanism (DAAM) of ZigBee, which considers residual energy and a network address to change the threshold criteria, making possible stable and energy-saving cluster set-up. This also ensures energy efficiency and increased network lifetime by uniforming energy distribution for the nodes. The secret is to calculate the number of clusters using an equation and exploit each node's network to adjust the CHs threshold in contrast to LEACH-C, EE-LEACH, and LEACH. The following results are obtained., Zhao et al., (2018).

Network Lifetime	**FND**	**HNA**	**LND**
LEACH	330s	450s	530s
LEACH-C	430s	500s	570s
EE-LEACH	390s	540s	500s
LEACH-M	500s	590	650s
Improvement ratio of LEACH-M	$\frac{16\%}{28\%}$	$\frac{18\%}{09\%}$	$\frac{14\%}{08\%}$

Table 2-1: Network Lifetime of LEACH and its Variants

Energy Consumption	**200s**	**350s**	**500s**
LEACH	52.70J	82.12J	99.43J
LEACH-C	39.73J	69.55J	95.65J
EE-LEACH	32.55J	62.33J	92.86J
LEACH-M	21.59J	48.15J	78.79J
Improvement ratio Of LEACH-M	$\frac{46\%}{34\%}$	$\frac{31\%}{23\%}$	$\frac{18\%}{15\%}$

Table 2-2: Energy consumption comparison of LEACH and its variants

The critical terms such as FND, HNA, and LNA are defined as follows.

- Stability period: This is the dead period of the first node in a WSN network, further known as the first node dead or first dead node (FND, FDN).
- Half-node dead (HND): the dead time of the half-sensor node is abbreviated as HND.
- Instability period: This is the dead time of the last sensor node in a network. The last node dead is another name for this (LND).
- Alive nodes: Overall, the number of member sensor nodes is up and running and functional and has not dissipated all of their energy.
- Residual energy: The available or remaining energy in total in the sensor nodes in a network, Saleem & Alabady, (2022).
- Network lifetime is the most significant time interval between First Node Death (FND) and Last Node Death (LND). Since the death and loss of a single sensor node and its data affect final results, more extended stability periods are a vital prerequisite for WSNs.

The next step to MODLEACH is improved MODIFIED LEACH (iMODLEACH), which considers three parameters, p, which is the probability of choosing a CH, s, (soft threshold), and h (hard threshold) and its impact on WSN performance. The value of p and h are iterated in MATLAB and the results obtained are as follows Ahmed et al., (2013).

S. No	***p***	***h***	**Max rounds in a network**	**The first dead node in a network**	**Ratio x = First dead node /Max rounds**
1	0.1	100	1095	160	0.146
2	0.1	2000	1248	148	0.123
3	0.1	300	1200	148	0.118
4	0.1	400	1261	131	0.104

5	0.1	500	1106	162	0.146
6	0.1	600	1216	159	0.131
7	0.1	700	1102	143	0.129
8	0.1	800	1207	133	0.110
9	0.2	100	1313	103	0.0784
10	0.2	200	1503	88	0.0565
11	0.2	300	1259	98	0.0778
12	0.2	400	1371	83	0.0605
13	0.2	500	1231	86	0.0698
14	0.2	600	1262	88	0.0697
15	0.2	700	1244	82	0.0565
16	0.2	800	1255	71	0.0431

Table 2-3: Probability and threshold variation effects on FND and rounds

Since p is a probability, it ranges from 0 to 1. h is a self-chosen value, either based on experimentation or typical threshold values, resulting in some unexpected results. The value of packets sent to BS gradually decreased from $h = 100$ to $h = 400$ then increased at $\mathrm{h} = 500$, and the same behaviour was observed from $h = 500$ to $h = 800$. As a result, a trade-off was detected between p and h. Additional experiments were performed with $p = 0.2$ and $h = 0.5$.

The two-level and three-level heterogeneous versions of LEACH are SEP and SEP-extended (SEP-E). The stability period, also known as the transition period, drives the concept of SEP, which was first put forth by, Smaragdakis *et al.,* (2004). A probability is formed for each node, called weighted action probability, that makes a node CH based on its remaining energy. The results are compared to LEACH, but direct transmission (nodes transmit directly to sink node, no clusters formation), and minimum transmission energy (transmission route is cheapest in energy cost).

Mohammad Hossein Homaei, (2022) describes the SEP method based on a few hypotheses. First, 1) heterogeneity; 2) The energy of each node varies; some, such as the CH or sink node, have higher energies; 3) randomness; 4) uniform distribution of these sensors; and 5) known coordinates for the sink and sensor. LEACH considers even energy distribution of sensor nodes; hence SEP exploits the node heterogeneity. The results indicate that SEP has a higher stability period and node throughput than energy distribution clustering protocols such as LEACH. The extended stability region results from some nodes' extra battery power/energy.

2.3 Nature-Inspired Algorithms For Optimal Coverage

Nature-inspired or Genetic Algorithms (GA) emerge as an exciting solution for the problem faced in WSNs. These algorithms provide optimal coverage, covering the issue of limited battery power of sensors and reducing the transmission of redundant information. Two approaches were contrasted in performance, namely, the combined Improved Genetic Algorithm and Binary Ant Colony Algorithm (IGABACA), and the second is Lion Optimization (LO) by Singh et al., (2021). The LO has shown a better lifetime of sensors and optimal coverage compared to IGABACA and hence better energy consumption metric, accurate sensing, and no transmission of redundant information.

When discussing multi-level clustering techniques to preserve energy consumption and im prove network lifetime for heterogeneous WSNs, a metaheuristic multilevel heterogeneous clustering technique (MMHCT) has been developed. To extend the network's lifespan, MMHCT lowers the energy consumption during cluster creation throughout network rounds and equally divides the load among the clusters to enhance the network lifespan. Multiple simulations against the latest techniques, such as Differential Evolution Based Clustering Routing Protocol (DEBCRP), are used to demonstrate the efficiency of the suggested clustering technique. The results show gains of 50%, 29.69%, 30.22%, 65.65%, & 37.17% in terms of network lifetime under 1-level, 2-level, 3-level, 4-level, & 5-level of energy heterogeneity, respectively Chaurasiya et al., (2021). This is a Genetic Algorithm using the concepts such as fitness function, crossover strategy and mutation and offspring production.

Another multi-hop routing protocol, inspired by Artificial Ants' feeding behaviour, is called Ant colony optimization-based WSN routing algorithm, as single-hop communications consume access energy for CHs. The ants have local information initially; however, they will release a substance named a pheromone while foraging the path for food hunt as ants or animals do in the wild, getting global optimization as well. This algorithm's energy consumption and average delay are compared to direct diffusion (DD), Samaras & Triantari, (2016). AOWSN (Ant optimization WSNs) accomplished the longer time and low energy consumption. Zhihui, (2015).

CHs in WSNs need to be appropriately load balanced because they use more energy and have more significant burdens than nodes because they are involved in more communication activities. The spider monkey-based optimization routing protocol highlights this. For antenna design and benchmark optimization projects, Spider Monkey Optimization (SMO), a naturalistic evolutionary approach that draws inspiration from the mechanism of how monkeys forage, is proven effective at solving the NP-hard problem of proper load-balancing-based clustering, Mittal et al., (2018).

Optimal selection of the nodes and forwarding packets in WSNs is crucial since wireless communication is prone to losses with distance, multi-path fading, and interference, bringing unreliability in WSNs. A simulation environment is created with sensor nodes spread in a rectangular grid of a 200m×600m, with uniform distribution for each node and node density set to 0.01, 0.009, 0.006, 0.005, respectively, and multi-hop reliability is simulated concerning nodes density and varying distance between the start and stop nodes. The higher density of nodes produced higher reliability, and the reliability of multi-hop dropped sharply with increasing distance between the start (source) and stop (destination) node, Song, (2013).

2.4 Game Theory based Algorithms in WSNs

The same problem of energy efficiency and network lifetime was attempted using a different multi-hop routing protocol, namely, game theory and coverage optimization (MRP-GTCO). The network lifetime is estimated from the node's death rate, and commonly found terms in literature are first dead node (FDN), last dead node (LDN), and half dead nodes (HDN). The paper mainly focused on the CH extra consumption compared to normal nodes and earlier dying characteristics. The authors Yao et al., (2022) demonstrated a comparative analysis of their proposed system with CROSS through MATLAB simulation, Koltsidas & Pavlidou, (2011), LGCA Xie et al., (2013), ECAGT protocols Q. Liu & Liu, (2017). Therefore, MATLAB simulation environment findings compare algorithms based on game theory.

Table 2-4: A Comparison of MRP-GTCO to CROSS, LGCA, and ECAGT (Yao et al., 2022).

Protocol	Clustering game and consideration factor	A factor of final CH selection	Multi-hop/one-hop
CROSS	Global competition, $p = 1 - w^{\frac{1}{N-1}}$	-------------------------------	One-hop routing
LGCA	Local game, $p = 1 - w^{\frac{1}{N_b - 1}}$	CSMA/CA Mechanism	One-hop routing
ECAGT	Local game, $p = 1 - w^{\frac{1}{N_b - 1}} * (\frac{Ei}{E_{ave}})^{\alpha}$	Residual Energy	One-hop routing
MRP-GTCO	Local game, $P_i = 1 - (\frac{\phi E_{CH} - E_{CM}}{\phi E_{CH}})^{\frac{1}{N_{bs_i} - 1}}$ Novel penalty coefficient ϕ	Residual Energy & CH position	multi-hop routing

When employing MRP-GTCO, the BS received more packets than its rivals in the 300m ×300m and 400m ×400m situations, as shown in Table 2-4; Thus, MRP-GTCO is shown to have a greater packet transmission rate and a good energy efficiency. MRP-GTCO also has higher live nodes, resulting in a higher energy utilization rate. Regarding network lifetime, the FDN, LDN, and HDN of MRP-GTCO are increased compared to additional procedures in table 2-1 due to the usage of the multi-hop routing technique.

The clustering algorithms of WSNs are created using a game theory approach. The game theory has been historically adopted as a powerful mathematical tool to ascertain the selfish and logical behaviour of creatures participating in the game, which are nodes in this case. These decisions are used to formulate the result of the game. This strategy is about accessing the profits and losses of each node, termed the Cost and Payment-based clustering Algorithm (CoPA). A weighted metric that combines each node's energy and transmission power regulates the alternation of CH nods. A Correlated Equilibrium (CE) algorithm was created using optimization theory. An adaptive regret matching (no-regret) strategy was employed to ensure that the probability distribution would eventually converge to the CE. The results of this algorithm are proposed for pure and mixed strategy Nash Equilibrium (MSNE) algorithms utilizing the nodes' efficiency and fairness as the main weight matrices. CoPA demonstrated better results for network lifetime and WSN system throughput, Attiah et al., (2017).

The duty cycle is an important concept that can be applied to clustering-based algorithms, resulting in Sleep-awake Energy Efficient Distributed (SEED) algorithm. The idea is to reduce the redundant information transmission to the BS to reduce energy consumption. However, there is an apparent disadvantage to this algorithm, which is idle listening problem. This is the period during which a radio of the sensor node is active; however, no transmission or reception occurs. SEED uses time division multiple access (TDMA), wherein a sensor is given a time slot (termed as a cluster member (CM) in SEED, and it has to turn on its radio despite having no data to transmit or receive. Hence, improvements were made in TDMA allocation to address the idle listening problem, and according to the traffic data rate and node energy, nodes were switched between sleep and awake mode. The solution is named energy and traffic aware sleep-awake (ETASA), and it achieved 15-16% improvements in the lifetime of sensor nodes compared to TEAR (traffic-energy aware routing) and SEED, Shagari et al., (2020).

In contrast to energy heterogeneity, traffic heterogeneity (high-traffic nodes consuming more energy) has received less study attention. This paper aims to make use of traffic heterogeneity for improved energy consumption. Table 2-5 summarizes the existing related works.

Table 2-5: Traffic rate and energy consumption Shagari et al., (2020).

Node	**Traffic rate (kb)**	**Initial Energy (J)**	**Energy consumed to send data to CH (J)**
Node A	10	1	0.5
Node B	7	0.7	0.2
Node C	6	0.6	0.15
Node D	8	0.8	0.3
Node E	2	0.5	0.02

The PEGASIS employs a chain-based approach, as depicted in Figure 2.2. A chain is formed using a greedy algorithm, in which the leader can be nearest to the BS or the sink node. Data aggregation can be implemented using token forwarding. The nodes nearest each other fuse or assemble their information and transmit it to the leader node, which then transfers it to the sink or BS. Data is only collected from individual nodes and sent to the sink node by the leader.; as a result, the leader node uses more energy and needs to have its energy usage adequately managed to increase energy efficiency, Lindsey & Raghavendra, (2002).

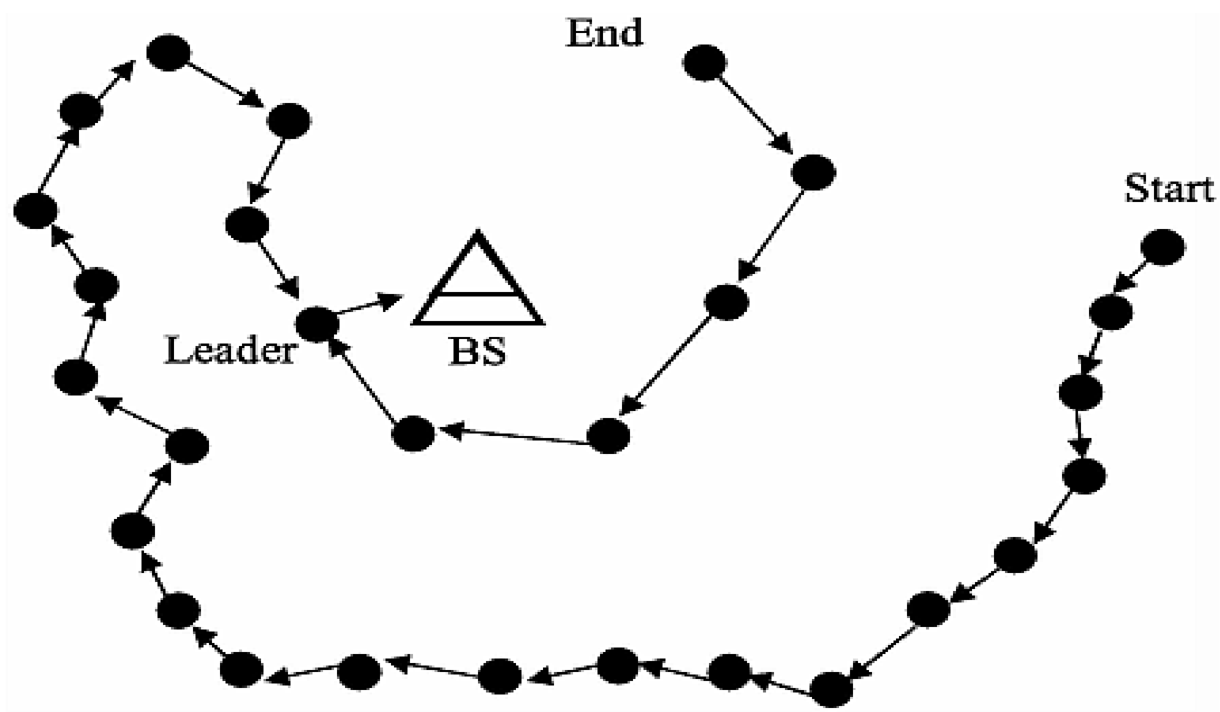

Figure 2-2: PEGASIS protocol and clustering mechanism

TEEN routing protocol is suitable for suddenly changing environments as it responds to changes in temperature, humidity, etc. It uses soft threshold (ST), and hard threshold (HT) concepts and CH selection in TEEN is similar to LEACH. HT is a pre-defined value, crossing which the node must transmit its data to CH. A node is turned on by ST principles based on the tiny difference to convey data to CH., Manjeshwar & Agrawal, (2001). The architecture of TEEN is illustrated in Figure 2.3.

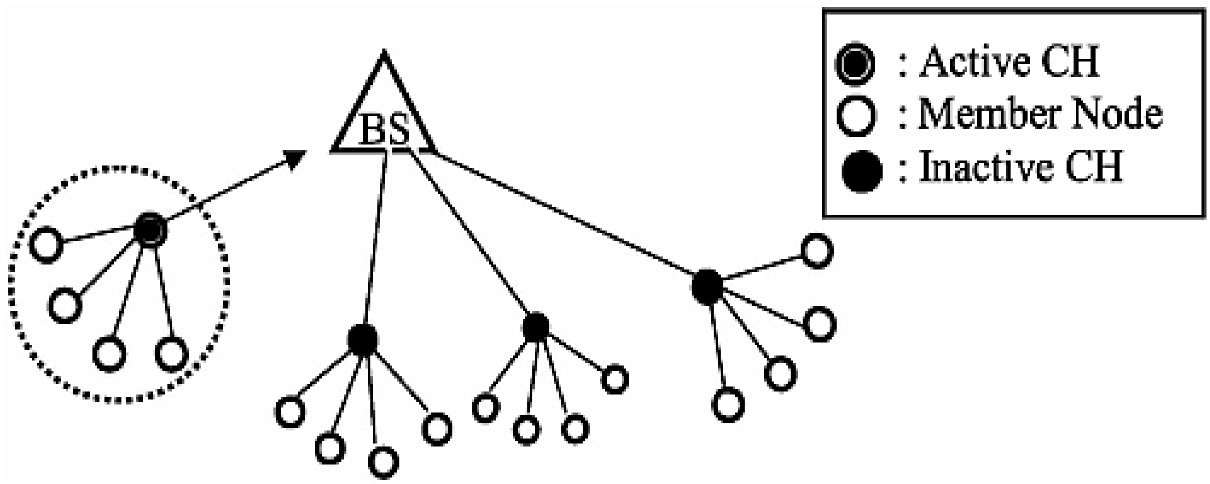

Figure 2-3: TEEN Nodes set-up

ETX (Expected Transmission Count): The number of MAC layer transmissions needed to successfully transmit a packet to a destination without any errors (including re-transmissions).

ETT (Expected Transmission Time): ETT should reduce with increased bandwidth for successful transmission.

Performance metrics	**LEACH**	**PEGASIS**	**TEEN**
Number of Hop	Small	Big	Small
ETX	Small	Big	Very Small
ETT	Small	Big	Very Small
Energy Dissipation	Small	Very small	Small

Table 2-6: Hop-count, ETX, ETT and energy consumption of TEEN, PEGASIS and LEACH

A comparison study of the effectiveness of LEACH, TEEN, and PEGASIS was performed by Khan et al., (2015). THE PEGASIS has a high hop count, ETT, and ETX, but has better energy consumption than other clustering forming protocols, with approximately 100-300% improvement. The threshold capability of the TEEN routing protocol incurs better energy consumption, as nodes are idle unless thresholds are crossed, leading to better longevity and throughput.

A MAC protocol combined with a routing algorithm that uses less energy (low power) is an intriguing strategy for maximizing energy savings in a WSN network. MAC protocols define the channel multiple access based on TDMA. MAC protocols have energy problems because of idle listening and scheduling node activities. The wireless channel's fading imposes another limitation as dependability problems worsen. The solution to this problem is provided by the channel quality-based metric. The energy-efficient unequal clustering (EECU), dynamic source routing (DSR), ad hoc on-demand distance vector (AODV), and multi-hop routing protocol with unequal clustering (MRPUC) are examples of low overhead protocols that are compared and recommended. The authors ultimately proposed a strategy for an inter-cluster routing protocol based on CQI and an intra-cluster MAC protocol with cross-layer communication capability, Sefuba & Walingo, (2018).

An optimization problem is solved while satisfying each node's energy restrictions by 1) creating a utility function and applying more energy to the relaying nodes and 2) calculating the exact amount of nodes needed in multi-hop WSNs to prevent energy gaps. Radio Frequency (RF) Energy harvesting is an efficient solution for increasing the network lifetime of WSNs by generating energy from ambient sources. Dedicated RF energy sources are deployed in the WSN field on purpose, which is used to compensate for energy holes. Energy holes arise as the sink has many-to-one communication, and chances for energy holes formation around sink nodes are very high. Once a hole is formed around the sink node, no more contact with the sink is possible, Lakshmi et al., (2018).

Fuzzy logic techniques increase the CH efficiency by making it adaptive, flexible, reconfigurable, and intelligent so that energy may be distributed among nodes sensibly for network scalability, lower energy consumption, and longer network lifetime. To handle the uncertainty in the wireless link model, type-2 fuzzy logic is applied to clustering, that is the reason why type-1 fuzzy logic was used by the majority of researchers in past research. Fuzzy logic addresses the varying percentages and uncertainties in the WSN by using type-1 fuzzy logic to type-N, as described in Nayak & Vathasavai, (2017).

According to the method of operation and type of target applications, routing strategies are categorized into the following three groups: hybrid, reactive, and proactive. Eight routing algorithms are covered in detail in this article by Bhattacharyya et al., (2010). LEACH, TEEN, Adaptive Periodic TEEN (APTEEN), PEGASIS, Sensor Protocols for Information via Negotiation (SPIN), Directed Diffusion (DD), Rumour Routing (RR), and Graphic and Energy Aware Routing (GEAR) are the eight protocols mentioned. The comparison results of these routing protocols are tabulated in Table 2-7.

Protocols	Power consumption	Network Timeperiod	Categorization	Multipath	Scalability
LEACH	Max	Excellent	Clustering	No	OK
TEEN	Max	Excellent	Clustering/reactive	No	OK
APTEEN	Max	Excellent	Hybrid	No	OK
PEGASIS	Max	Excellent	Clustering/reactive	No	OK
SPIN	Restricted	OK	Proactive/flat	Yes	Restricted
DD	Restricted	OK	Proactive/Flat	Yes	Restricted
GEAR	Restricted	OK	Location	No	Restricted

Table 2-7: Multiple Algorithms Comparison Based On Homogenous Clustering

The transmission time and throughput are the two most essential parameters encountered in WSN. The three algorithms, direct transmission, LEACH, and EEE-LEACH, are compared concerning throughput and transmission speed. EEE-LEACH is two level clustering formation algorithm. CHs are formed in the first layer of the algorithm; CHs acquire information from sensor nodes, fuse it, and then Master CHs (MCHs) are included in the in-2nd clustering layer algorithm. By measuring the distance between them and combining the data from various sensors, the CH looks for the closest MCHs. MCHs receive data from the nearest CHs, aggregate it, and compress it to send it to the nearest BS. The threshold probabilities for each CH and MCH are used to calculate the number of MCHs, which is algorithmically lower than the CHs. As shown in Figures 2-4 and 2-5, compared to direct transmission and the LEACH protocol, this improves EEE-energy LEACH's efficiency and lengthens the network lifetime. However, the two-layer clustering technique extends the transmission time and reduces throughput.

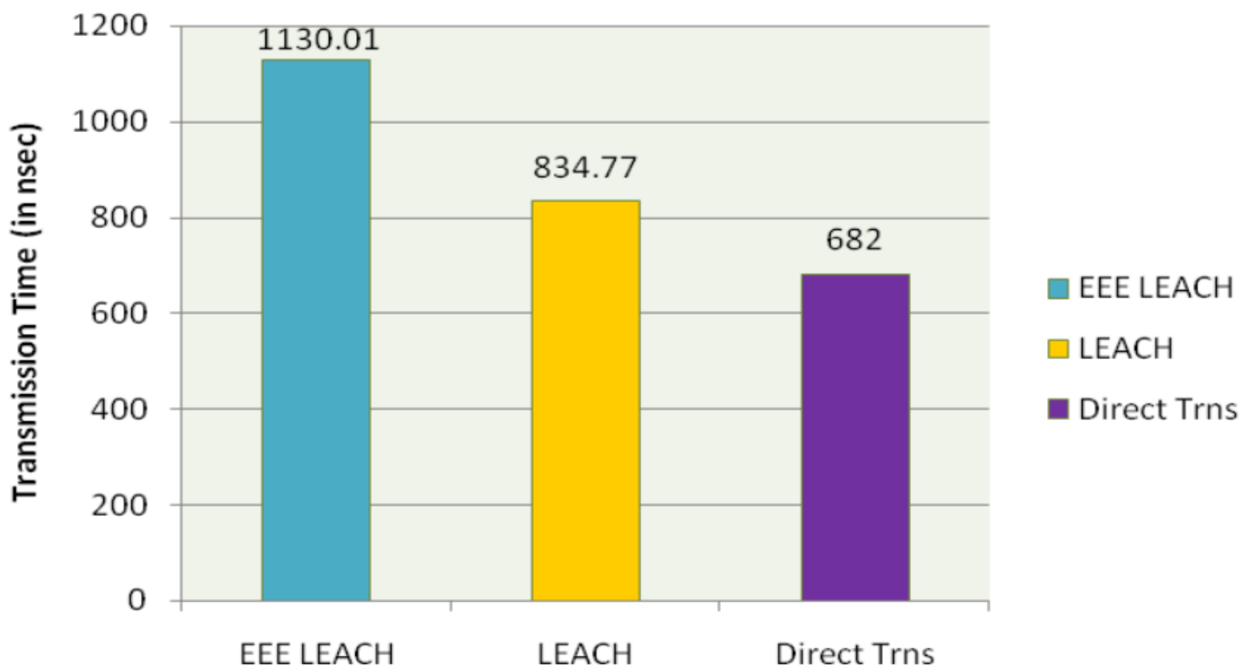

Figure 2-4: Transmission time comparison of EEE LEACH, LEACH and Direct transmission

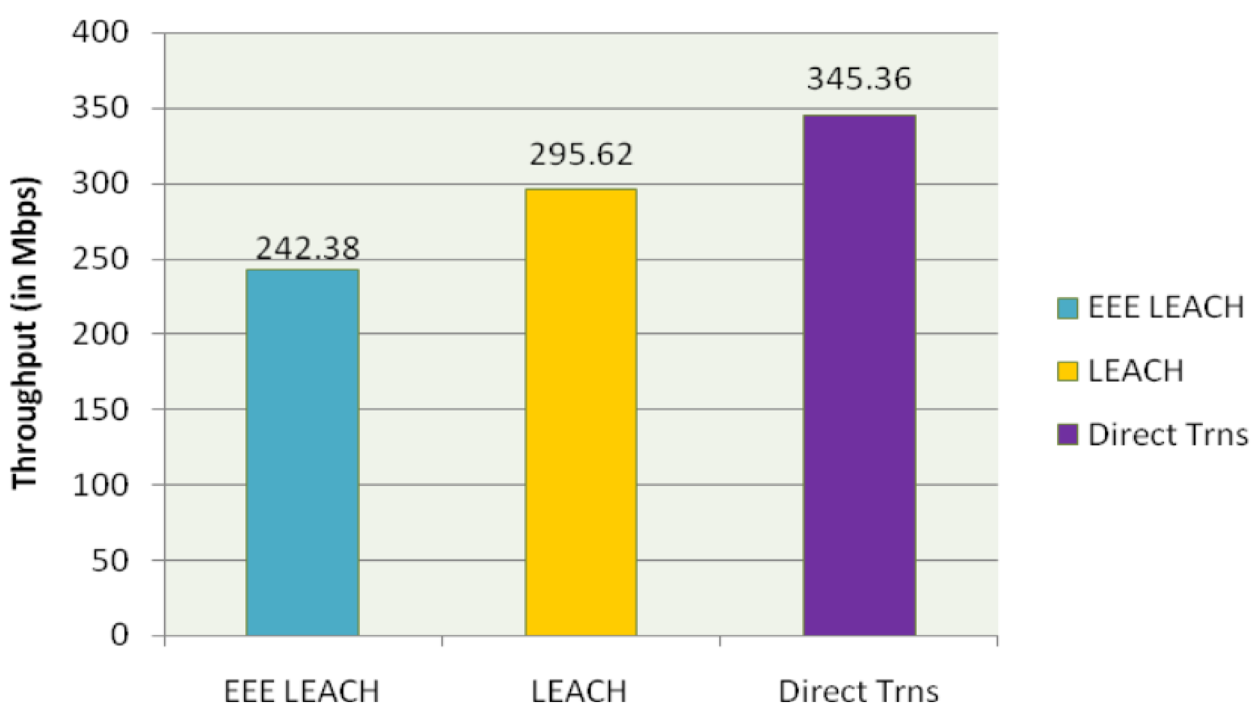

Figure 2-5: Throughput comparison of EEE LEACH, LEACH and Direct transmission

2.5 LEACH method

The standard LEACH method distributes the energy in the nodes equally. The set-up phase and the steady state phase are the two phases that each round of Leach's system operation consists of. One round starts from one clustering phase to the next stage of CH selection. In LEACH algorithms, the ideal amount of rounds can be determined by N/K, where N stands for nodes, while K is an approximated number of clusters.

❖ <u>Set-up Phase</u>

During the setup stage, LEACH decides the CHs based on an election procedure. A probability value is chosen as a basis for CH selection algorithms with the condition that there are k predicted clusters every round. During the selection process, A random value between 0 and 1 is generated by nodes and contrasted with an arbitrary threshold. The node is regarded as the cluster leader if its random value is lower than the threshold.

$$T(n) = \begin{cases} \dfrac{P}{1 - P(rmod\ \left(\frac{1}{P}\right))} & if\ n \epsilon G \\ 0 & Otherwise \end{cases} \tag{1}$$

In equation (14), P is the probability value chosen such that the estimated numeral of CHs is K for every round, and G is the number of CHs that were denied the chance to become CHs in 1/P repetitions. The r is one for round 0, and after becoming cluster chiefs, nodes cannot revert to being CHs for the subsequent 1/P operations. The CH can send a message to all nodes that are members after becoming CH following the threshold and random value generation comparison procedure. Members receive that advertisement based on signal strength. Each member node is assigned a TDMA slot for transferring their data to the CH, and the transceiver can be turned off unless there is time to send the data. The CH can transport data collected from member nodes to the BS in an aggregated form, Uday B. Mavdiya & Limbad, (2015).

In the next phase of CHs and clustering formation, the member nodes based on signal strength communicate to the CH. The information is exchanged in a "join packet" that contains their CSMA identifying IDs in order to join a certain cluster.This provides CH information with regard to a cluster's number of nodes and the respective IDs. After receiving packets inside of the cluster, the CH sends the TDMA slots information to each member node. The CH then creates a

TDMA assignment table and a random CSMA code before broadcasting these to each member node. The steady-state phase then begins..

❖ Steady State Phase

The data transmission phase is during this time when the nodes can turn off their radios and send their data to the CHs. To receive data from the sensor nodes in the cluster, the CHs radio must be turned on. Abidi & Ezzedine, (2017).

❖ Radio Interference

The communication lines between the various clusters interfere with one another as the clusters are produced. Several clusters employ CDMA (code division multiple access) to prevent this interference. Each CH is given a code for the cluster, and each member node receives this code via broadcast. As a result, each cluster's individual coding can be cracked to obtain the data. The use of CDMA lessens interference. The radio signal from the nearby clusters is simply filtered and does not taint transmission on the relevant channel, Ali Qasim Alrubaye, (2022).

2.6 Procedure:

❖ Advertisement

The appropriateness of each node as a CH is established by comparing its random number, which ranges from 0 to 1, to a pre-defined value.

The first loop is this one. When a CH is chosen, messages are sent throughout the entire network, informing it of the node's characteristics.

❖ Cluster set-up

This phase consists of the choice of which nodes will be connected to which by calculating the distance between each node and the CH. The sensor members of the cluster choose which clusters to join based on this distance calculation. The energy usage obviously decreases with decreasing distance and vice versa.

❖ CH transfer

The CH energy is compared to a threshold that has been established at this point. If the CH energy does not reach the threshold, the algorithms automatically promote the next node in the clusters with the highest energy to the position of CH. Higher energy nodes are only considered as CHs, and nodes with lower energy and rapid death rates are automatically avoided., Lu et al., (2017).

2.7 The limitations of LEACH

The limitations of the LEACH protocol are as follows.

1. The distributed probabilistic method is used for a node to be selected as CH, while the non-cluster nodes join the clusters based on signal strength. This ensures low data overhead, but the CHs are not guaranteed to be uniformly distributed over the entire network. The load imbalance means a CHs is not located at the end compared to nearby the sensor, reducing network lifetime.
2. Another assumption is the similar energy distribution and equal chance to be elected for the CH. This energy and probability distribution uniformity are impractical since most non-cluster nodes join the clusters based on signal intensity, whereas the distributed probabilistic technique is utilized to choose a node as CH. Sensor nodes are spread in the field heterogeneously. Hence, LEACH needs to be modified for heterogeneity.
3. The source node is required for direct data transmission to CHs. If the CH and source nodes are far, there will be more energy consumption for far-located sensors. LEACH then assumes that CHs will use a single-hop link to send the data that has been aggregated to the sink node. This single-hop transmission is expensive in cases where the sink is far from the CHs.
4. Then, using TDMA, each node is allotted a time slot, and CHs notify the nodes when they can transmit data.
5. Another incorrect assumption in LEACH is that all nodes conserve the power to transfer their energy to the sink node, which might not be accurate for some of the energy-constrained nodes, Arumugam & Ponnuchamy, (2015).

Considering the limitations of the LEACH protocol, the agenda of this report has been kept to come up with a LEACH variant or any novel algorithm that can be used to improve the previously existing algorithms, and improvements are made. LEACH and its derivative, such as MODLEACH, have been compared using simulations, as shown in the results section.

2.8 Proposed Method (LEACH Modification)

In this chapter, improvements to the current approaches are suggested. The field's size, node count, probabilities, and a few other factors, such as energies, have all been altered, and observations have been made regarding how the changes have influenced the performance.

In Table 2-8, a comparative study of different LEACH variants is conducted, and their performance is compared concerning varying parameters. This comparison of performance parameters has been extracted from the following research papers, Fanian et al., (2016; Kandpal et al., (2015; Maurya & Kaur, (2016):

Algorithm Name	Hoping Scheme	Improvements over LEACH	Results of Improvements
LEACH -C	Multi-hop	With the known location of member sensors, a centralized clustering strategy	Reduced data transmission cost and increased lifetime
TB-LEACH	Multi-hop	The fixed period for CH selection	Network lifetime improvement
TL-LEACH	Multi-hop	Two-level hierarchy of CHs	Improved scalability and throughput
ED-LEACH	------------	For CH nomination, remaining energy with euclidean distance was used.	Better CH distribution, network lifetime
MR-LEACH	Multi-hop	Multi-hop routing and -hierarchy of CHs.	Better network lifetime
LEACH-GA	Multi-hop	Genetic algorithm used	Improved network lifetime
LEACH-V	intra-cluster communication	Adoption of vice CH and increased lifetime proposed	Prolonged lifetime
Cell-LEACH	Cells are formed instead of clusters	cell head and CH formation and cluster division in cells	Even/uniform clustering method and improved network lifetime
Improved V-LEACH	------------	In order to choose vice CH, consider the maximum residual energy, minimum energy, and minimal separation distance.	Better Network Lifetime

LEACH-MAE	------------	For moving sensor nodes, improved CH selection.	Uniform energy distribution and enhanced network lifetime
LEACH-SCH	------------	CH selection to reduce clustering overheads.	Prolonged network lifetime
EEM-LEACH	------------	CH used to find the multi-hop path	Enhanced network lifetime
LEACH-CKM	------------	MTE transfers data and K-means categorization utilized	Enhanced coverage and better network lifetime
Op-LEACH	------------	Free TDMA slots for data transmission	Improved throughput and network lifetime
(LEACH)^2	------------	With multiple sink nodes, the Sensing region is split into four areas.	Enhanced throughput and network lifetime
LEACH-M	Single-Hop	No location and mobility awareness	------------
LEACH-ME	Single-Hop	location and mobility awareness	------------
LEACH-B	Single-Hop	No location and mobility awareness	------------
LEACH-P	Single-Hop	No location and mobility awareness	------------
W-LEACH	Single-hop	No location and mobility awareness	------------
N-LEACH	Single-hop	No location but Mobility Awareness	------------
LEACH-MF	Multi-hop	No location but with mobility awareness	------------
FL-LEACH	Single-hop	No location and mobility awareness	------------
LEACH-R	Multi-hop	No location, with mobility awareness	
WLEACH (Wise LEACH)	Single-hop	No location but with mobility awareness	
Improved FZ-LEACH	Multi-hop	No location, with mobility awareness	
LEACH (uneven clustering)	Multi-hop	No location, with mobility awareness	
EL-LEACH	Single-hop	No location and mobility awareness	

Ad-LEACH	Single-hop	No location and mobility awareness
K-LEACH	Single-hop	No location but with mobility awareness
DD-LEACH	Multi-hop	No location and mobility awareness
Q-LEACH	Multi-hop	No location and mobility awareness
DAO-LEACH	Multi-hop	No location and mobility awareness

Table 2-8: Tabular Comparison of LEACH Variants, Covering Most Of The Literature.

The impact of mobility on the performance of WSNs and specifically energy consumption and network lifespan, has been studied in many research studies. Through comparing the simulation results for four scenarios—the event is stationary, moves randomly, moves with a simple four-path, boids path is used for the traversing of the event—the behaviour of WSNs for the environment when the event moves with a specific movement path is explored. The results indicate that for the case where the event is travelling arbitrarily has the weakest performance. With more sensor nodes present, the goodput characteristic falls. When the Tr is less than 10 pps for the boids model, the output is unstable. With an increase in Tr, then the dissipated energy rises. The poorest of the four possibilities, according to simulation data, is random movement's energy use. The boids model's energy consumption is the lowest of the four. This demonstrates how energy consumption in large-scale WSNs can be reduced by event mobility using a boids model., Yang et al., (2010).

Boids is a model for simulating the flocking behavior of birds in a WSN (Wireless Sensor Network). The model was first introduced by Craig Reynolds in 1986 and is based on three simple rules: separation, alignment, and cohesion. These rules define how individual "boids" (representing sensors in a WSN) interact with one another to produce the overall behavior of the flock.

In a WSN context, the boids model can be used to simulate the behavior of nodes in a network, for example, to study the formation of clusters, the distribution of nodes, or the performance of routing algorithms. By modeling the interactions between nodes, the boids model can help to evaluate and optimize the design of WSNs .

PPS stands for Packets Per Second and represents the number of packets transmitted by a device in a given time frame, usually one second. It is a measure of the network's transmission rate.

Throughput Rate (Tr) is a measure of the amount of data successfully transmitted from a source to a destination in a given time period. It's typically measured in bits or bytes per second.

Goodput is a measure of the effective utilization of the available bandwidth, taking into account any losses due to errors, retransmissions, and other overhead. Goodput is calculated as the ratio of the amount of useful data transmitted to the amount of data that was transmitted, including overhead. In WSN, goodput is a crucial metric because it helps to determine the efficiency of the network and to optimize it for the application's requirements.

Two mobility patterns (optimised spiral and Gaussian mobility patterns) are created for base station repositioning to maximise WSN lifetime and present a quantitative analysis of the proposed mobility patterns in comparison to well-known mobility patterns in a novel optimization framework to investigate the effects of sink mobility patterns on WSN lifespan (random and grid mobility patterns)., Cayirpunar et al., (2015) . These two studies, despite being pertinent, have not highlighted the impact of sink mobility on service quality indicators. In order to evaluate the effects of sink mobility patterns on WSN lifetime, a unique optimization methodology is used in this study. For base station repositioning to enhance WSN lifetime, a few mobility patterns (optimal spiral and Gaussian mobility patterns) are presented, along with a quantitative analysis of the proposed mobility patterns in contrast to well-known mobility patterns (random and grid mobility patterns). In order to characterise the effect of mobility patterns on WSN lifetime, we actually combined an exact mathematical programming model (MIP) with a heuristic search space (the mobility patterns), Yagouta, Jabberi, et al., (2017).

Identifying the best configurations that provide the best trade-off between energy conservation and QoS metrics is essential. To study this trade-off, three sink mobility models Random Walk (RW), Random Waypoint (RWP), and Gauss Markov (GM), are studied. Simulation figures provide the insight that to node density of 0.08 node/m2 (800 nodes in 100m 100m) with a packet rate of 1 packet/s, or for a packet rate of 4 packets/s for each sensor node with node density of 0.01 node/m2, reasonable and acceptable numbers can be achieved to characterize the trade-off between the two parameters. Beyond these final node density and packet rate thresholds, the network uses more energy, and the other quality of service (QoS) metrics rapidly decrease., Yagouta, Gantassi, et al., (2017).

The mobility-based routing protocol (MBC) is another way to improve CH selection and reduce energy consumption. This work makes two contributions: First, a mix of average node energy and average node speed is used to create a centralised clustering algorithm that regularly chooses an ideal set of CHs. This differs significantly from MBC, which selects CHs based on the node residual energy and the current node speed using a distributed clustering algorithm. Second, certain elements are considered before a detached node joins an ideal cluster. A centralised energy-efficient clustering routing protocol for mobile nodes (CEECR) is suggested to reduce energy loss and boost the packet delivery ratio or throughput. With less mobility and more energy, this protocol uses a central control algorithm to build a superior set of CHs.

A detached node's best CH is chosen based on the total weights. In terms of average energy dissipation and packet delivery ratio, simulation findings show that CEECR performs better than its competitors and is more energy-efficient overall. In future development, it has been investigated into how to create the best-distributed protocol for CEECR. For instance, the best CH will be selected using the position forecast of MNs., Zhang & Yan, (2019).

Wireless sensor networks (WSNs) frequently employ low-capacity, non-rechargeable batteries.

Creating an energy-efficient routing protocol generally significantly affects network lifespan extension. To improve the energy efficiency of WSNs, the novel distributed 2-hop cluster-routing protocol (D2CRP) is introduced in this study Chen et al., (2022). Each node in the 2-hop range gathers information about its neighbours during the cluster-building phase to spread the 2-hop cluster fully. The energy-efficient cluster head (CH) in each 2-hop cluster is chosen by considering both the transmission distance and residual energy. Each member node can send packets to the CH or its 1-hop neighbour after the CH has been generated. Several chains can be built between CHs via their neighbouring CHs closer to the BS to decrease the overall transmission distance for intercluster communication to BS. Packet transfer happens through energy-efficient intracluster and intercluster routing. To minimize energy consumption for both intracluster and intercluster communications, the ideal cluster number for 2-hop clustering is also derived. In terms of network lifetime, energy consumption, and packet transmission, simulation results show that the optimal cluster number of D2CRP can be reached and that it effectively outperforms the other four state-of-the-art competitive routing protocols LEACH, RLEACH, PEGASIS, and two-tier distributed fuzzy logic-based protocol (TTDFP).

2.9 Gaps in Literature Review

Each WSNs protocol comes with various restrictions; this implies that not a single algorithm fits all scenarios, and the majority of the literature that is currently available has ignored one of the below-provided points

- When choosing CHs, the majority of researchers have ignored taking into account the separation between the BS and the sensor node.
- The DEEC and LEACH versions miscalculated the correct number of clusters in every cycle.

2.10 Summary of literature review

In the literature review, many routing protocols are examined. The evaluation of the literature on clustering procedures for WSN starts using the direct transmission approach, which involves sending sensor data connecting nodes to BS directy without the formation of cluster or CHs. The research study gradually worked toward better-performing clustering protocols, such as LEACH and its variations. The HEED, PEGASIS, and TEEN clustering protocol versions are also reviewed. Power consumption and network longevity were the significant characteristics to be analyzed and enhanced in this study; however, these methods' performance was examined using a variety of other factors, such as the number of dead nodes and alive nodes vs the number of rounds.

The nature-inspired and game theory-based algorithms are also reviewed with a systematic performance analysis comparison of the game theory-based and nature-inspired algorithms.

CHAPTER 3

3. System Modelling

While understanding the problem of wireless channel modelling of a simple system consisting of a wireless channel, a transmitter, and a receiver has always been challenging. The study attempts to achieve this in this chapter with the help of channel modelling and propagation theory in multi-hop WSNs.

After reviewing the literature, two approaches are drawn out for this problem; the first is an experimental and measurement-based approach for channel modelling, path loss computation, etc. Another choice is to use a theoretical model derived from the most recent academic research to represent the system and channel. The best and most obvious option would be to develop a new or innovative WSN routing protocol.

3.1 Radio channel model

The fundamental challenge in modelling the energy consumption of a WSN is to determine how much energy is needed to send x bits of data from the transmitter side to the reception side. One transmitter (Tx|), one receiver (Rx), and a channel make up this basic radio channel model of the first order, as depicted in Figure 3.1.

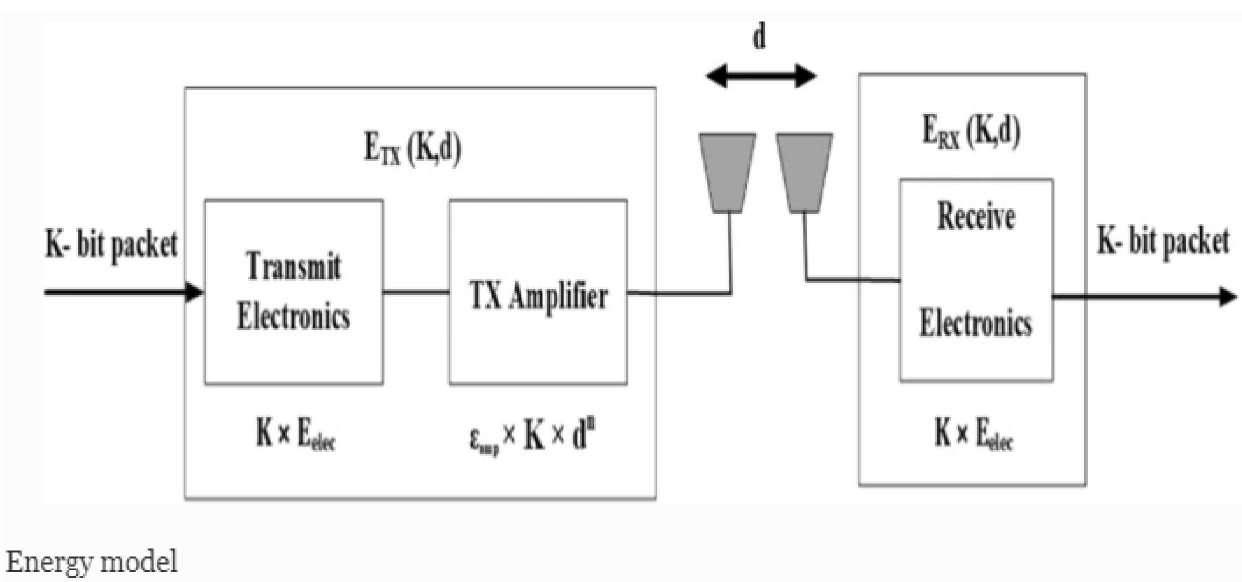

Figure 3-1: Radio channel modelling

The amount of transmitter energy needed to send a group of n bits from a transmitter to a receiver at d miles apart will be determined by the equation below.

$$E_{TX}\,(n, d) = E_{tc}(n) + E_{amp}(n, d) = n.E_{trans} + n. \;\; \in_{amp}. d^{\alpha} \tag{1}$$

According to the literature, the open space and the multi-path models are two different energy consumption models considered in this research. The total energy used to send a packet of k bits is provided by (5), and the energy consumption at the receiver is provided by (8):

In both of the equations mentioned above, do is the distance threshold that the below formula can calculate:

$$d_o = \sqrt{\frac{\in_{fs}}{\in_{mp}}} \tag{2}$$

where the $\in_{fs}$ is the energy of the open loss model and $\in_{mp}$ is the energy of the multi-path model. These energies, which correspond to the multi-path fading model and the free-space model, respectively, are the amplification energies at the transmitter for both models, Xin Qu, (2012).

For the open loss model, the energy consumption per bit of the transmitted bits is given by:

$$E = \in_{fs} d^2 \tag{3}$$

For the open multi-path model, the energy consumption per bit of the transmitted bits is given by:

$$E = \in_{mp} d^4 \tag{4}$$

The overall transmitter energy dissipation can be given by:

$$E_{TX}(k, d) = \begin{cases} E_{elec} * K + K * \in_{fs} * d^2 & d \le d_0 \\ E_{elec} * K + K * \in_{fs} * d^4 & d > d_0 \end{cases} \tag{5}$$

The communications channel is considered to be symmetric, and the energy dissipated by a sensor node in the transmission of k bits per packet can be approximated by the following equation:

$$E_{TX}(k,d) = E_{TX\_elec}(k) + E_{TX_{amp}}(k,d) \quad (6)$$

Finally, the energy dissipation in the receiver is given by:

$$E_{RX} = E_{ELEC} * K \quad (7)$$

Just like the equation for transmission of energy, the received energy for a sensor node in receiving k bits /packets is approximated by:

$$E_{RX}(k) = E_{RX_{elec}}k + k\, E_{ELEC} \quad (8)$$

E_{elec} is the amount of energy lost per bit on the receiver side. Every packet of bits is comprised of data and overhead or redundant bits, with data bits providing the information and overhead/redundant bits giving extra information, such as modulation and coding techniques for trustworthy data transmission, encryption techniques for privacy and security, and source and destination addresses. Regardless of the size of the packet, the overhead is a constant size. Hence, the overhead may utilize minimal spectrum and energy for big packet sizes, which would improve WSN spectral and energy efficiency. Additionally, the throughput is raised.

N_{data}, $N_{overhead,}$ $E_{overhead}$ N_{data}represent the number of data bits, how many overhead bits there are, the energy of overhead bits, as well as the energy of data bits, respectively, Behera et al., (2018). This makes the total energy of the packet as given by:

$$E_{packet} = N_{data}N_{overhead}E_{data}E_{overhead} \quad (9)$$

Where the efficiency is given by:

$$\eta = \frac{N_{data}\, E_{data}}{E_{packet}} * 100 \quad (10)$$

3.2 Power Consumption Modeling:

The power consumed by the node varies according to distance. The modelling transceivers in WSNs cannot be performed without the knowledge of different parameters that effects the energy consumption in WSNs. The power consumption model is essential for an accurate radio energy model. The power consumed increases in direct proportion to distance, as shown in the Figure 3-2.

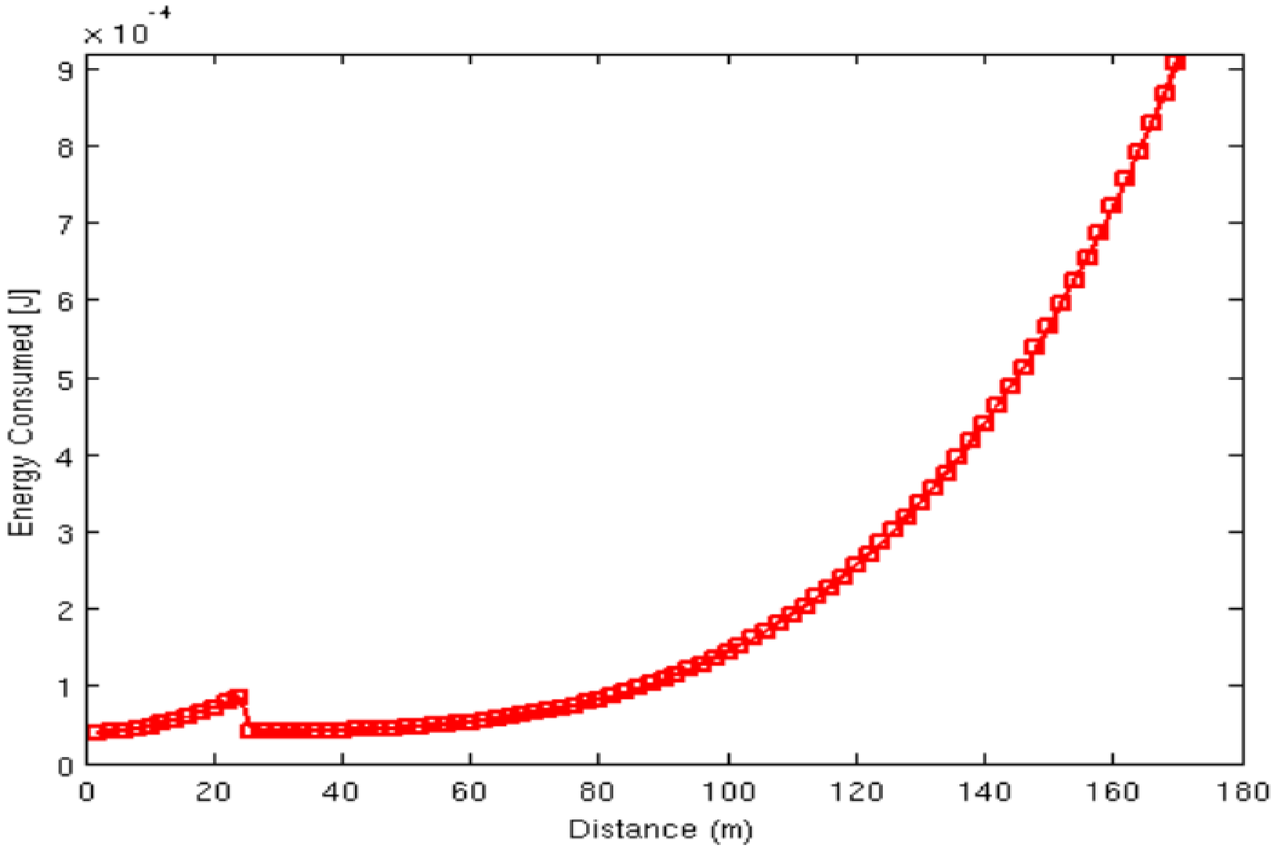

Figure 3-2: Energy consumption plotted w.r.t to distance

hence to transmit an n-bit packet beyond a distance d, the energy dissipated by the radio is given by:

$$E_{TX} = \begin{cases} nE_{elec}\, l \in_{fs} d^2 & d < d_{crossover} \\ nE_{elec}\, l \in_{PE} d^4 & d \geq d_{crossover} \end{cases} \qquad (11)$$

The parameter $\in_{fs}$ is based on the free space model, and this parameter can be approximated on the PE model. Any distance below the crossover distance is considered a short distance. For that case, the energy dissipation is according to the square of the distance, according to the FSL model. Power dissipation is determined by the fourth power of distance, while any distance above the crossover distance is in the PE model.

$$E_{TX} = \begin{cases} nE_{elec}\, l \in_{fs} d^2 & d < d_{crossover} \\ nE_{elec}\, l \in_{PE} d^4 & d \geq d_{crossover} \end{cases} \qquad (12)$$

The power transmitted is attenuated using the FSL model when the distance between TX and Rx is less than the crossover distance and using the two-ray ground plane propagation model when the crossover distance is more than d, Kamarudin *et al.*, (2016). The crossover distance can be calculated by:

$$d_{crossover} = \frac{4\pi\sqrt{L}h_{Tx}h_{RX}}{\lambda} \tag{13}$$

where L is the system losses, h_{Tx} and h_{RX} d transmitting and receiving antenna heights, and λ is the wavelength.

The power level required for successful data transmission and reception on the receiver is adjusted based on the distance as given by the equation:

$$P_{TX} = \begin{cases} \alpha_1 \, P_{sensitivity} d^2 & d < d_{crossover} \\ \alpha_1 \, P_{sensitivity} d^4 & d \geq d_{crossover} \end{cases} \tag{14}$$

Where $\alpha_1 = \frac{4\pi^2}{G_{TX}G_{RX}d^4}$ and $\alpha_2 = \frac{1}{G_{TX}G_{RX}h_{TX}h_{RX}}$ and $P_{sensitivity}$ is the receiver sensitivity level.

3.3 Simulation Parameters

❖ Initial Conditions

Table 3-1 provides the initial parameter settings for the simulations; adjustments are performed in accordance with desired results from the algorithm. However, adjustments can be made based on performance gains, an algorithm's demands for better rsults, or the parametric limitations of the simulation set-up.

Symbol	Parameter	Unit	Magnitude
S	Size of the field (sensors deployment field)	m	100*100
N	Number of nodes		100, 200,300
p	likelihood of being chosen as cluster leader		0.1, 0.2,0.5

ϵ_{fs}	The transmitter energy for the free space model	J(Joule)	10e-11
ϵ_{mp}	The transmitter energy in multi-path mode	J(Joule)	10e-15
E, ini	each node's starting energy	J(Joule)	0.5, 0.7
TR. E	Transmit or receive the energy of each sensor device	J(Joule)	5*10-8
PS	The packet size for data and control packets	b (# of bits)	4000
R	No of rounds		1000, 10000

Table 3-1:Initial parameters for the simulation

CHAPTER 4

4. Results Comparison and Discussions

The results chapter discusses the simulation method and results obtained for the typical and most common WSNs routing algorithms. MATLAB simulation environment is used for the study.

The study begins with the LEACH protocol as an example, works with the LEACH variations, compares the results from LEACH and MODLEACH, and evaluates the power usage and network lifetime. Many other WSN algorithms are also analyzed, including DEEC and its variants comparison in terms of critical parameters, as well as the outcomes of several academic researchers' research on DEEC and its variants.

Finally, DEEC and its variants are simulated for parameters performance comparison. LECAH and MODLEACH are simulated, and compared as well. Other WSN algorithms, such as TEEN and PEGASIS, SEP, and TSEP, will be simulated depending on time availability. Performance enhancement comparison and contrast have been made for the conventional methods and their improved versions.

4.1 DEEC and its Variants:

The first clustering algorithm to be studied is the DDEC algorithm and its variants. DDEC is the first clustering algorithm that makes clusters and chooses a CH for a specified number of nodes. These nodes communicate (transport data) with the CH, which communicates with the BS. Every node shares data on energy consumption and network lifespan and CH is selected based on an arbitrary probability threshold. Since the research focuses on improvements, the term scaling factor is improvised, meaning that nodes might be capable of communicating with the BS directly. This change should enforce intra-cluster communication at more minor power levels, ensuring reduced power consumption. This means the amplification power saves 10% for a 10*10 field size. Redefining threshold probability and utilizing neighbourhood information are two more changes that can be made in addition to this scaling factor redefinition to enhance performance, Yadav A. & Kumar S.,(2016). The IDEEC protocol is suggested as per the below specifications.

Table 4-1: WSN model parameters, field size, and number of nodes

Parameter	Value	Parameter	Value
Network Field	100m x 100m	E_{DA}	5 nJ/bit/Message
Number of Nodes	100	d_o	70m
E_{elec}	5 nJ/bit	Message Size	4000 bits
E_{fs}	10 pJ/bit/m^2	P_{opt}	0.1
E_{amp}	0.0013 pJ/bit/m^2	$Et_{hreshold}$	$E_0/4$
E_0	0.5J	Scale Fcator	10

The results in Figure 4-1 indicate that IDEEC performance surpasses the DEEC and LEACH by a significant margin (approximately from 2-4 times for rounds ranging from 1000-4000) in terms of the overall quantity of packets delivered to BS.

The same improvement in performance is obtained for IDEEC than DEEC and LEACH. The dead nodes stopped at 60 for maximum of 4000 rounds (number of iterations as on the X-axis) for IDEEC, while for DEEC and LEACH, the dead nodes crossed 100 for a maximum of 3000-4000

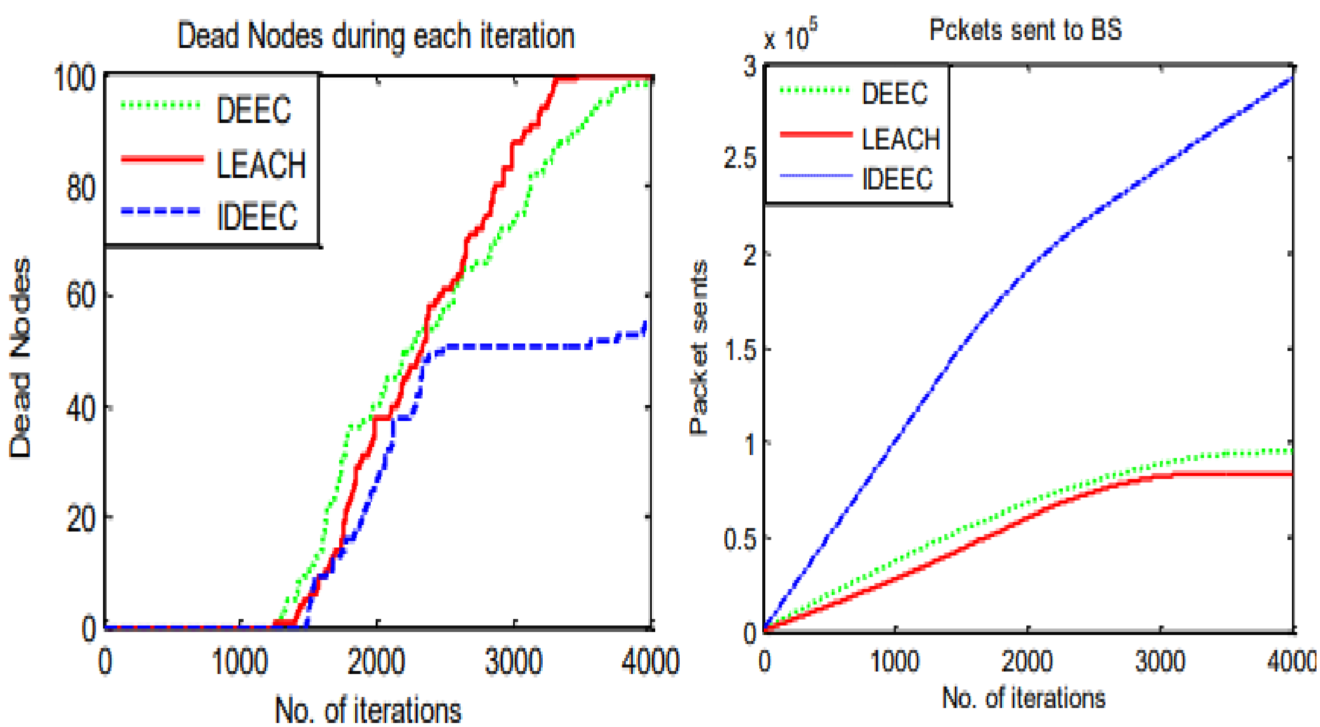

Figure 4-1: DEEC, LEACH comparison to IDEEC in Dead Nodes and dispatched packages to BS.

iterations or rounds.

In the next simulation set-up, the following DEEC variant algorithms are simulated and compared for analysis, DEEC, developed-DEEC (DDEEC), Enhanced-DEEC (EDEEC), and Threshold-DEEC (TDEEC). The SEP protocols consider two levels by design, normal and advanced nodes, to create multi-levels. However, DEEC is three-level clustering based on regular, advanced and super nodes. In terms of stability period, the three-level heterogeneous routing protocols for WSNs have significant energy level consumption disparities across regular, advanced, and super nodes. DEEC and DDEEC perform well, Qureshi et al., (2012). The following four figures are obtained for simulations from the TDEEC proposed method and its performance is merited as opposed to DEEC, DDEEC, and EDEEC. The TDEEC performs better than DEEC, DDEEC, and EDEEC regarding the number of alive or operational nodes for a specified number of rounds (maximum limit 10000). While the improvement is modest until 4000 rounds, the TDEEC outperforms other algorithms after 4000-10000 rounds with many alive nodes (from 5-40) during those terminal rounds.

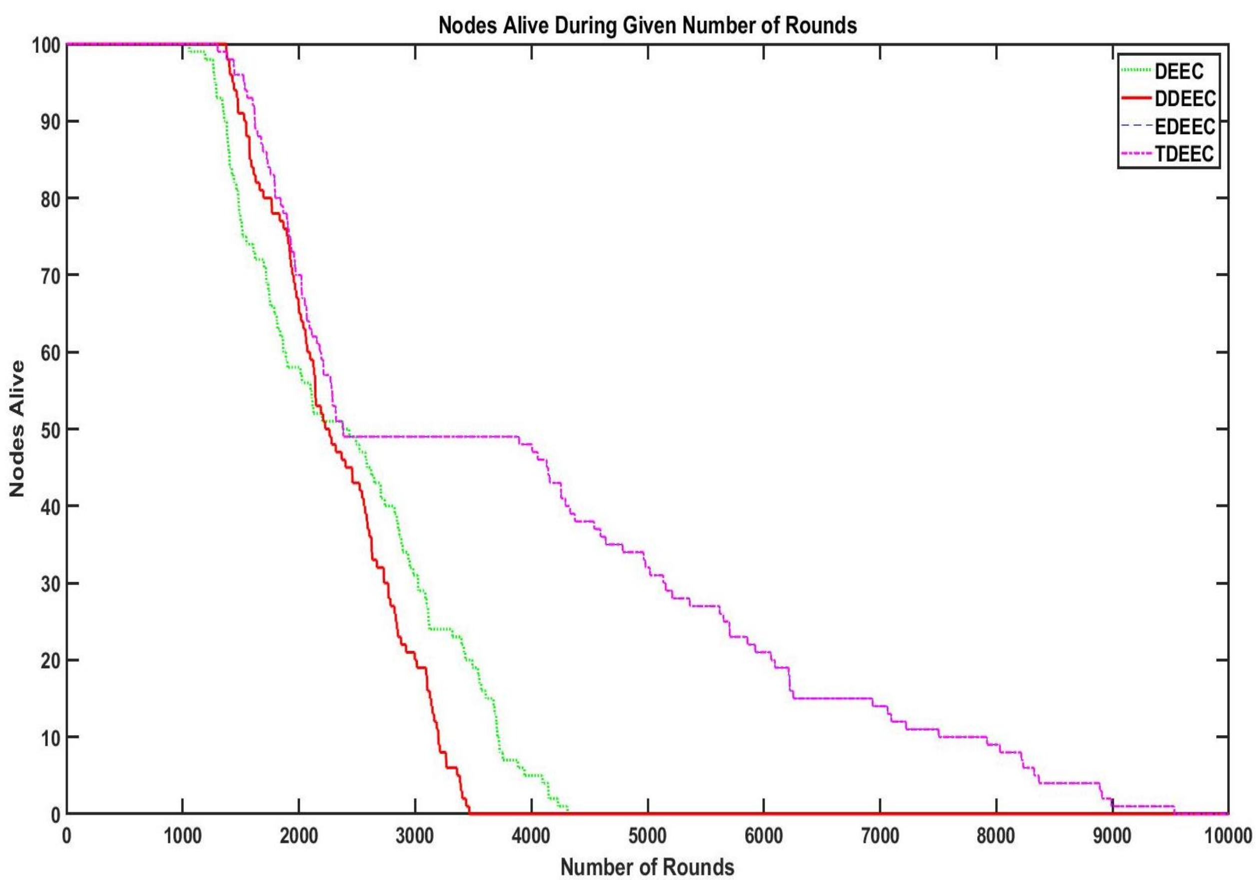

Figure 4-2: DEEC, DDEEC, EDEEC Comparison to TDEEC (Proposed Methodology) in Alive Nodes.

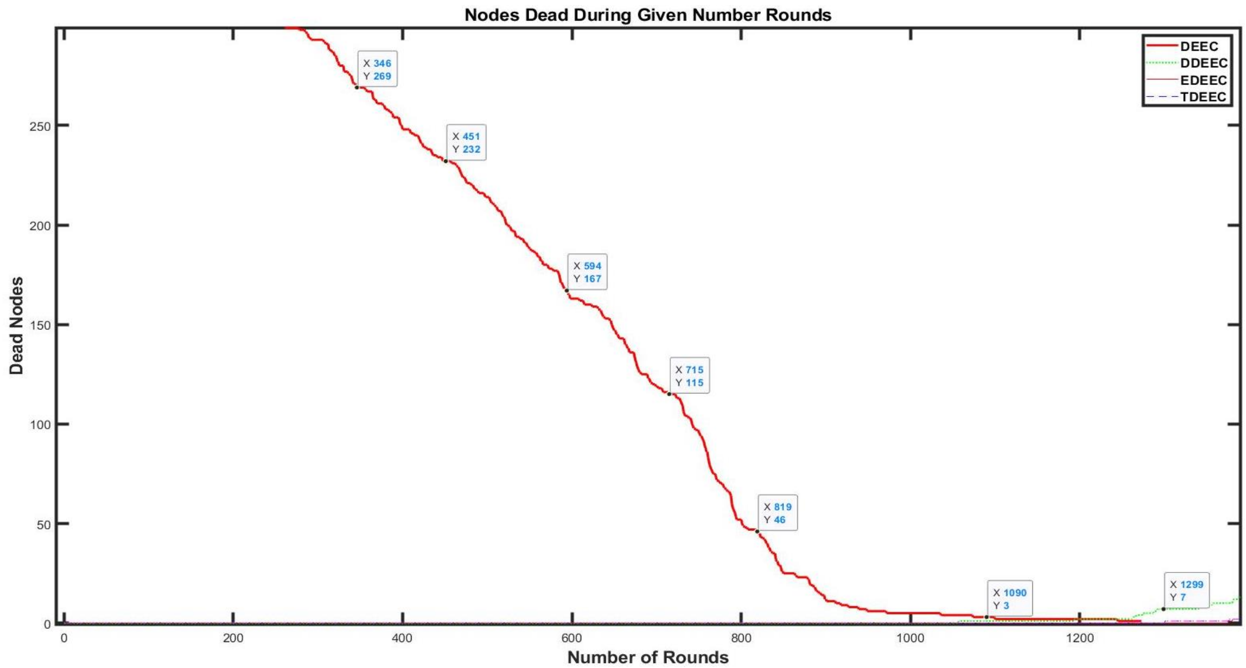

Figure 4-3: Dead Nodes Performance Comparison of TDEEC, with DEEC, DDEEC, EDEEC

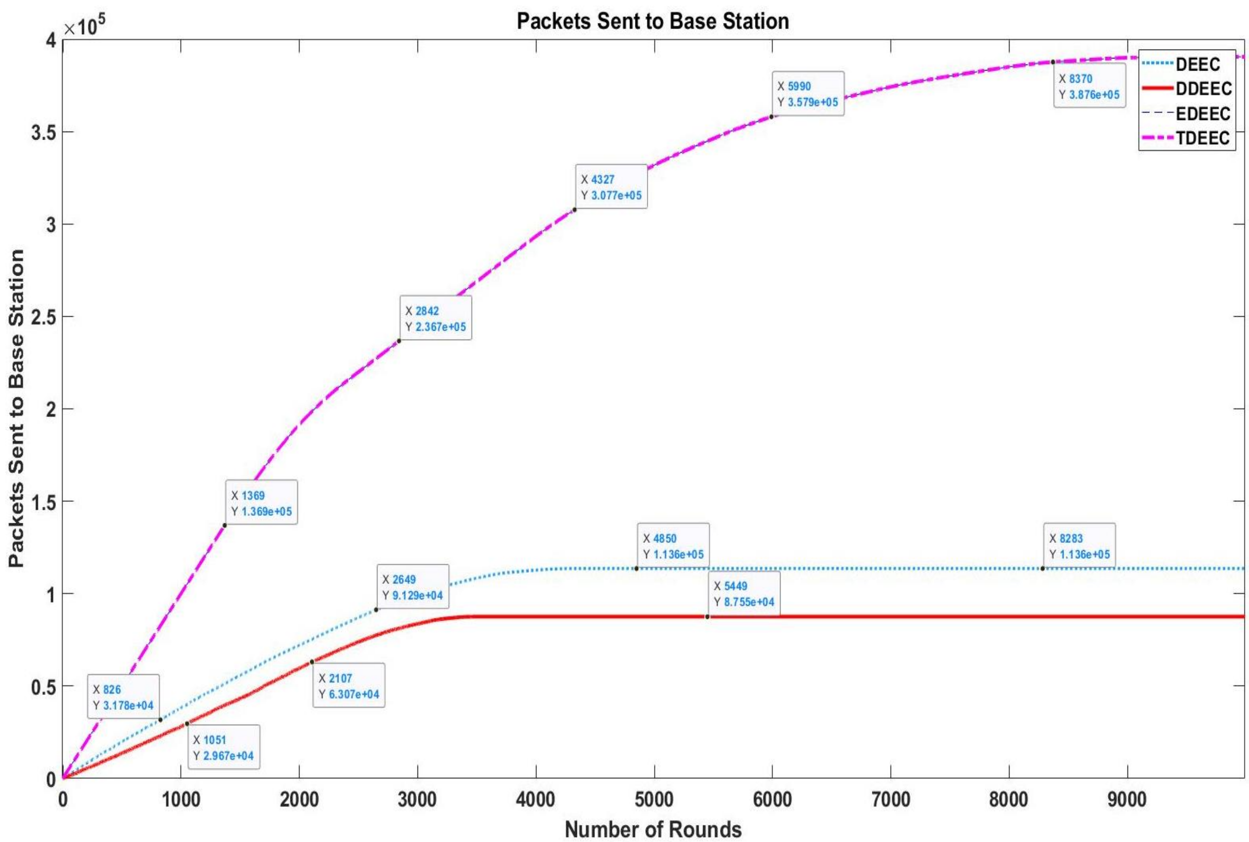

Figure 4-4: Packets sent to BS, Performance Comparison of TDEEC, with DEEC, DDEEC, EDEEC.

In terms of the quantity of dead nodes, the performance of the suggested algorithm (TDEEC) is not very impressive when compared to other algorithms in this study.

Concerning packets sent to BS, the TDEEC outperforms the DEEC, DDEEC, and EDEEC, by a factor of almost 2-to-4. This indicates a better algorithm design compared to previous methods. The CH formation is also better for TDEEC than DEEC, DDEEC, and EDEEC, although not clearly conceivable from Figure 4-5.

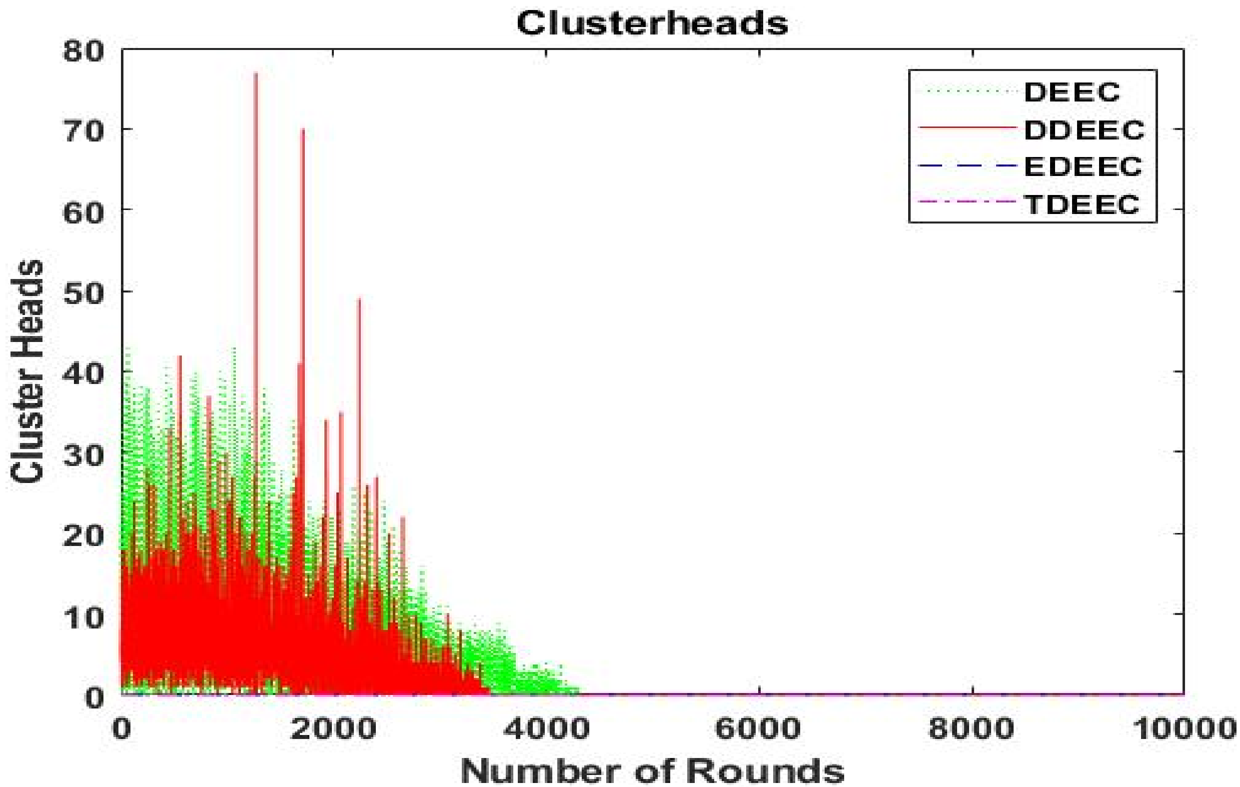

Figure 4-5: CHs Formation Performance Comparison of TDEEC, with DEEC, DDEEC, EDEEC

EDEEC and TDEEC perform poorly in terms of network endurance when compared to the other two. In contrast, EDEEC and TDEEC display good lifespan performance and stability in multi- and three-level heterogeneous WSNs with slight energy level variations between normal, advanced, and supernodes. The data demonstrate that the proposed algorithm, TDDEC, outperforms DEEC, DDEEC, and EDEEC in terms of all four parameters, i.e., greater performance in terms of operational (alive) nodes, more packets transmitted to the BS (almost four times), and fewer dead nodes in a set number of rounds.

4.2 LEACH

The LEACH is the pioneer of clustering-based algorithms. An area of 300*300 m^2 and 200 nodes distributed in that area is used as the testbed. The comparison is drawn between the LEACH and MODLEACH algorithms based on some essential parameters, such as dead nodes, packets sent to BS and CH, and alive nodes.

4.3 Modelling WSN in MATLAB for LEACH

For LEACH, 200 nodes are Spread randomly in a 300*300 m^2 open field. The following parameters are used as the remaining variables in the algorithm for the WSN network during the 3000 rounds. The modelled WSN in MATLAB looks like Figure 4-6 after the sensors are randomly distributed across the field.

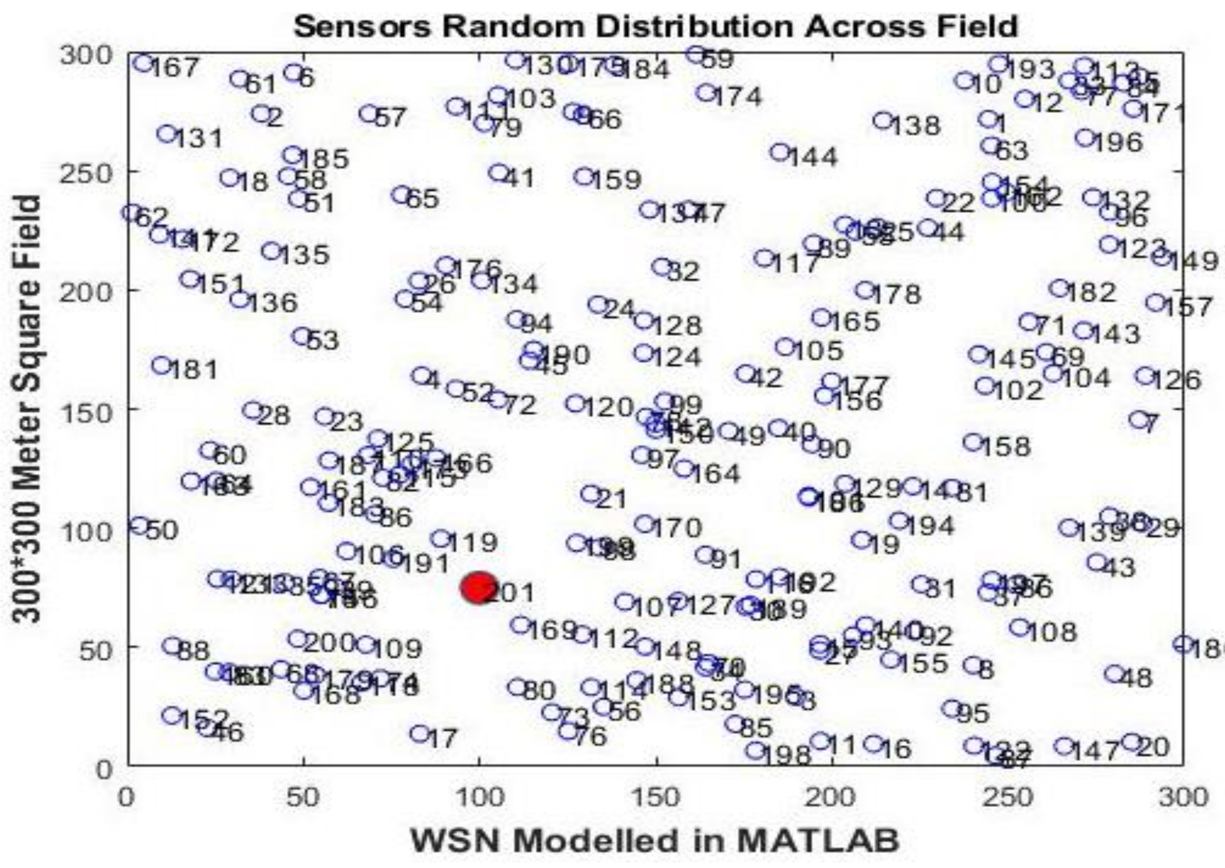

Figure 4-6: WSN modelling in MATLAB (Sensors Spread Randomly In Given Field)

The clusters are formed as per the hold decided by the LEACH algorithm. The following Figure 4-7 illustrates the cluster formation for LEACH.

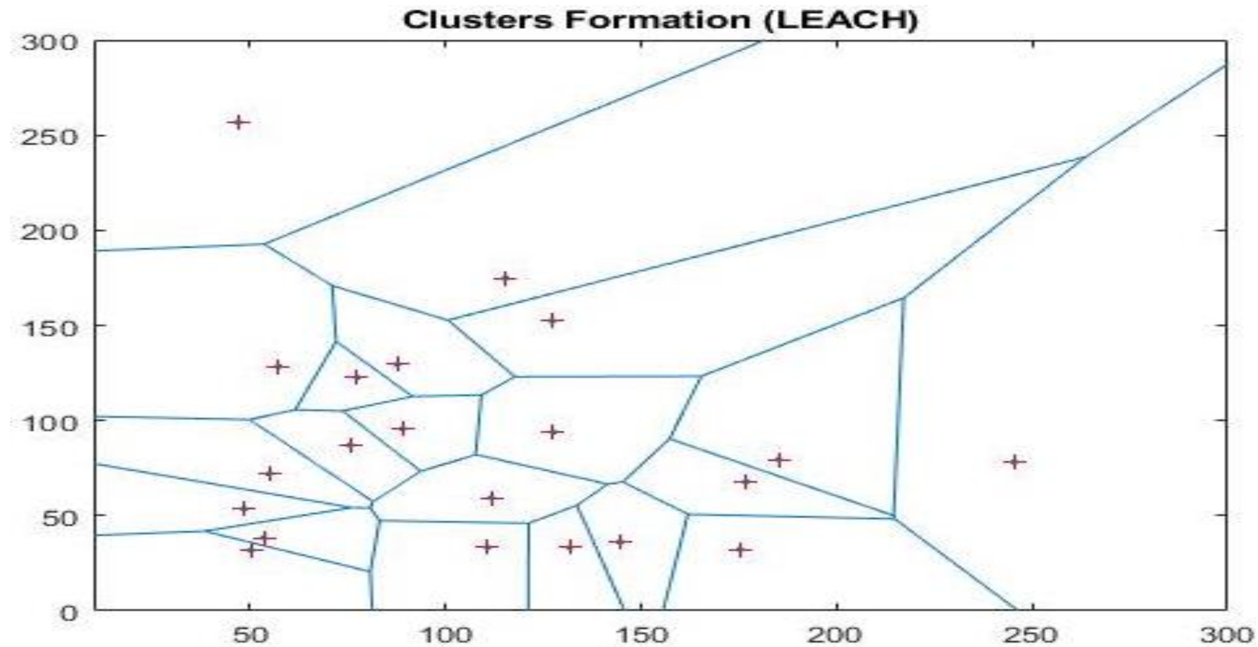

Figure 4-7: Clusters Formation In LEACH

The average residual energy for each node is depicted in the following figure. The average energy for sensor nodes is rapidly in inverse proportion to the increasing number of rounds. After 1000 rounds, the energy has reached a constant of zero joules level throughout the 3000 rounds.

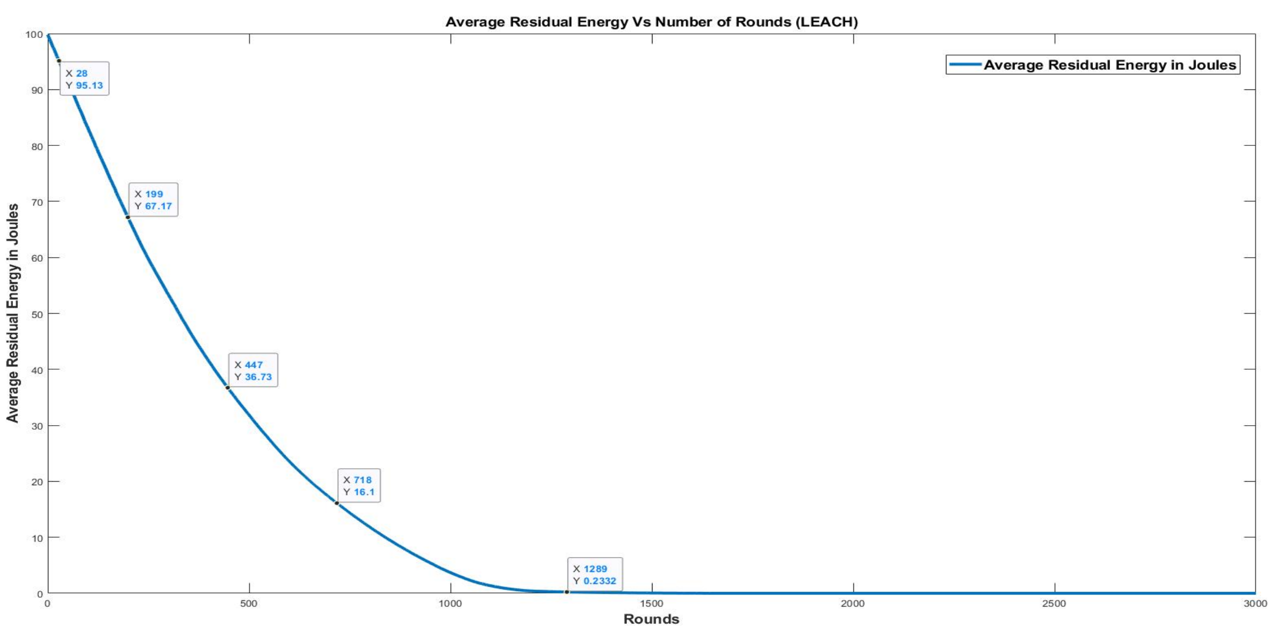

Figure 4-8: Average Throughput for LEACH against Number of Rounds

The next described is the throughput obtained for the LEACH WSN model. The following equation gives the formula for throughput.

$$\text{Throughput} = (\text{Size of the packet} / \text{Transmission time}) \quad (15)$$

The throughput increases linearly according to the number of rounds, and it is observed after 1000 rounds, it reached a maximum value and remained constant until the erst of 2000 rounds.

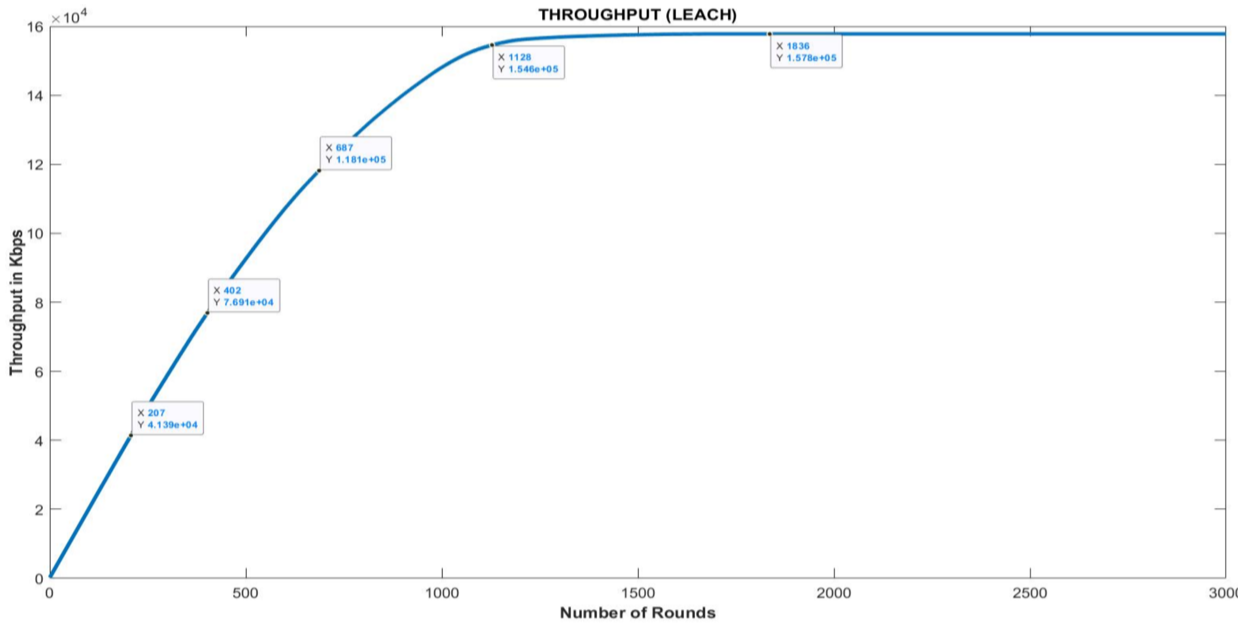

Figure 4-9: Maximum Throughput for LEACH vs Number of Rounds

4.4 Modelling WSN in MATLAB for MODLEACH

The LEACH CH election algorithm can be enhanced based on the threshold mechanism. Soft and hard thresholds, two types of thresholds, are used in MODLEACH. MODLEACH tends to lower the amount of energy required by the network by efficiently replacing CHs after the first round and employing two distinct transmitting strengths for communication between clusters and between CHs and BSs. In MODLEACH, a CH will only be replaced if its energy drops below a predetermined threshold, lowering the routing load. Because of this, the CH replacement process involves consuming the CHs remaining energy at the start of each round. Additionally, MODLEACH is used to build soft and hard thresholds to contrast the efficiency of these protocols in throughput and energy usage, Mahmood et al., (2013).

For MODLEACH, 200 nodes are spread randomly in a 300*300 m^2 open field. The parameters used are the same LEACH and the remaining variables in the MODLEACH algorithm for the WSN network during the 3000 rounds. The subsequent WSN modelling is obtained after the random distribution of sensors in a given field. The sensors distribution is less congested in MODEACH than in LEACH sensors distribution.

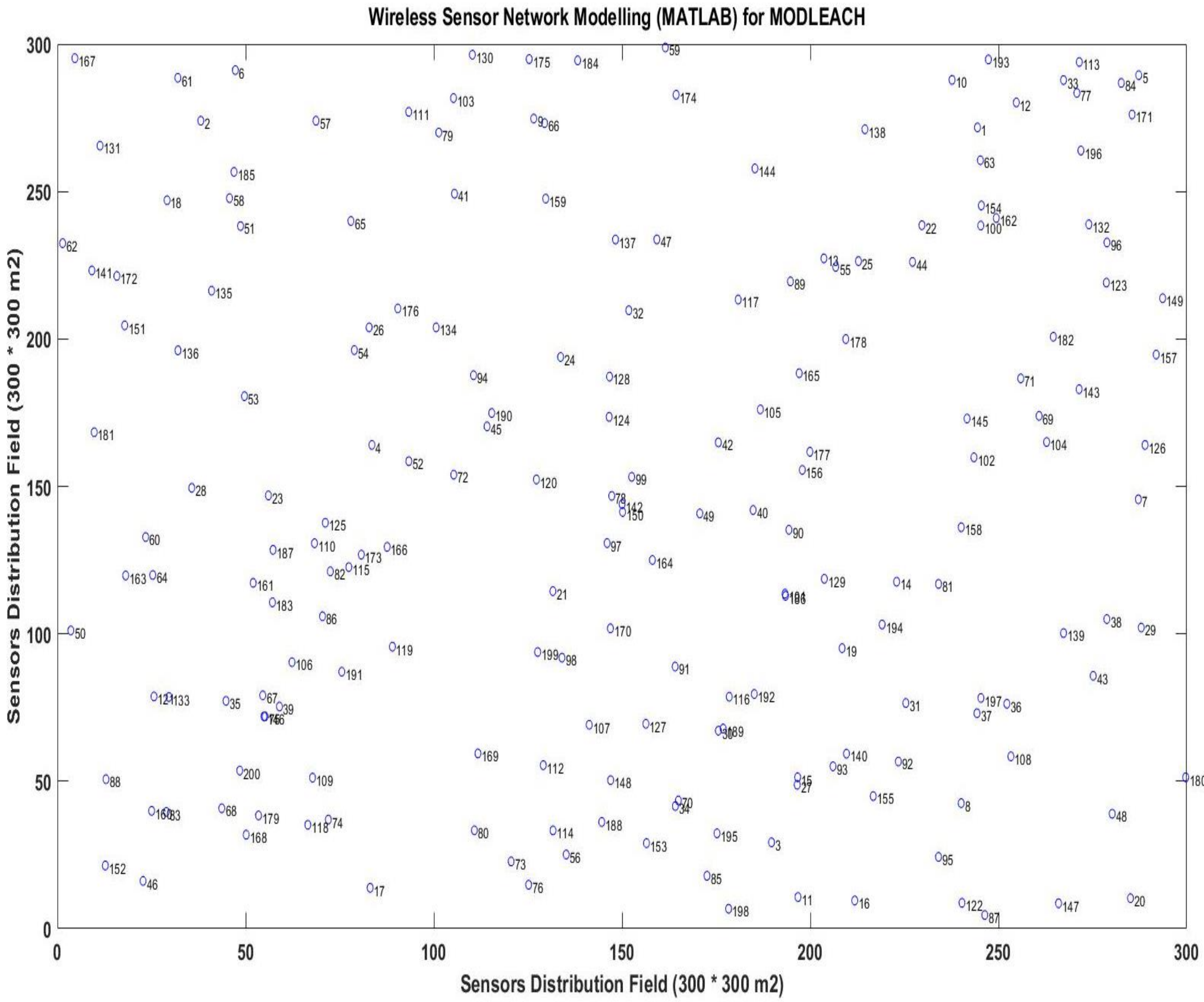

Figure 4-10: WSN Modelling in MATLAB for MODLEACH (sensors spread randomly in given field):

In the following section, the results of LEACH and MODLEACH are compared against the parameters such as alive nodes, dead nodes and packets sent to BS and CH.

4.5 The results comparison between LEACH and MODLEACH:

The results from MODLEACH and LEACH are compared based on essential parameters such as dead nodes, alive nodes, packets sent to BS, packets sent to CH, and the number of CHs formed during the first clustering head set-up phase.

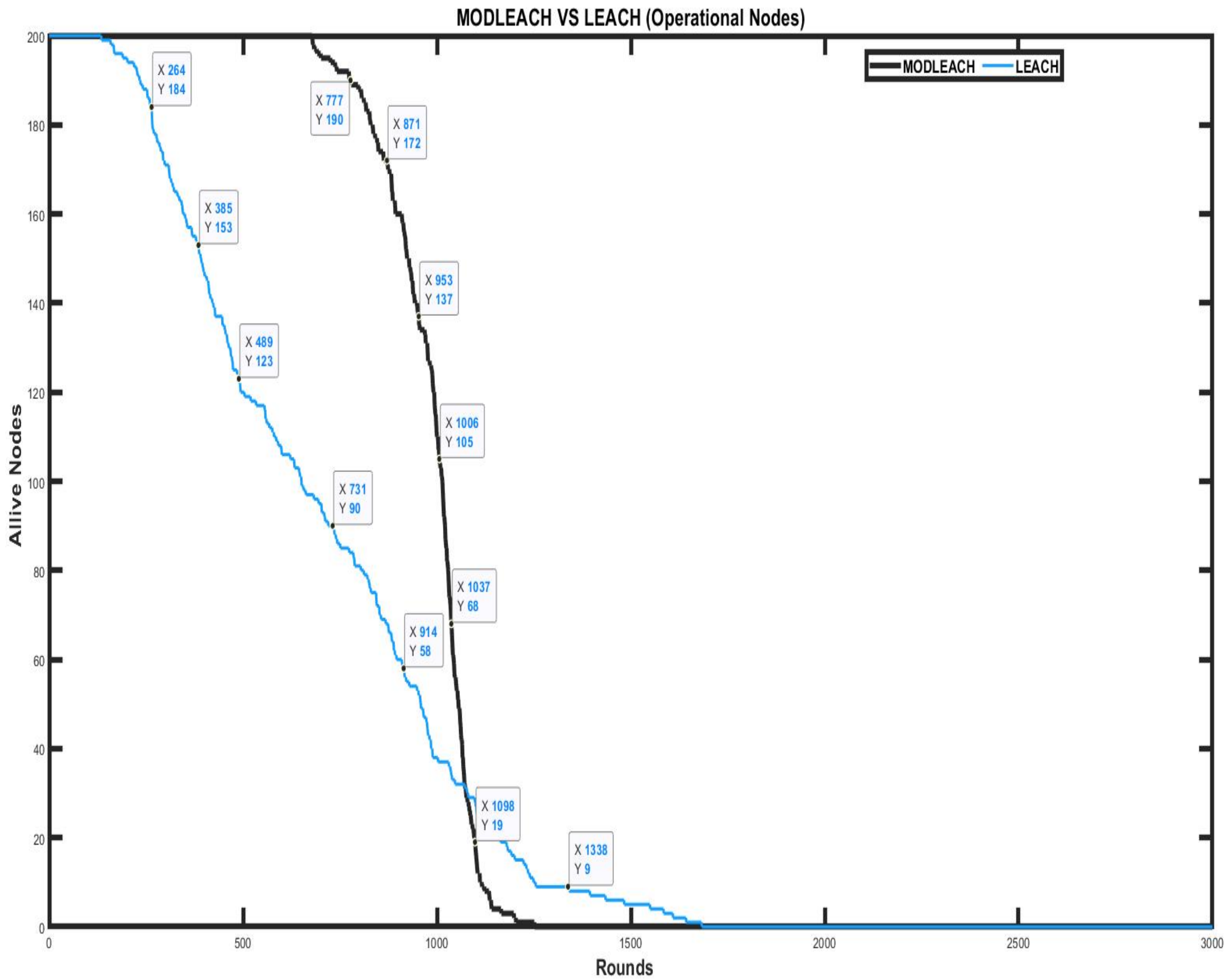

Figure 4-11: Functional (Alive) Node's comparison of LEACH vs MODLEACH:

Based on Figure 4-11, in the first 500-100 rounds, MODLEACH's operational nodes remained above 200, whereas LEACH's functional nodes dropped from 200-10 between 300-1500 rounds. MODLEACH has demonstrated enhanced performance in this study.

As illustrated in Figure 4-12, after 500 rounds, there are twice as many dead nodes in LEACH compared to MODLEACH; MODLEACH greatly enhanced the network lifetime and improved energy consumption by a factor of two.

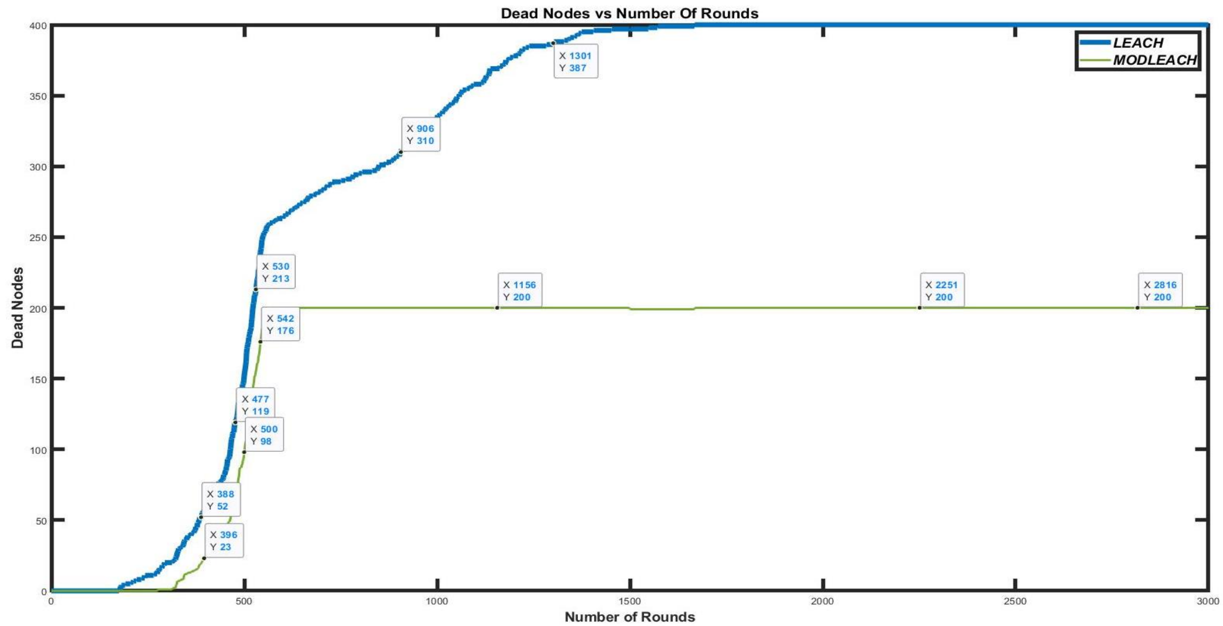

Figure 4-12: Dead Nodes comparison of LEACH vs MODLEACH.

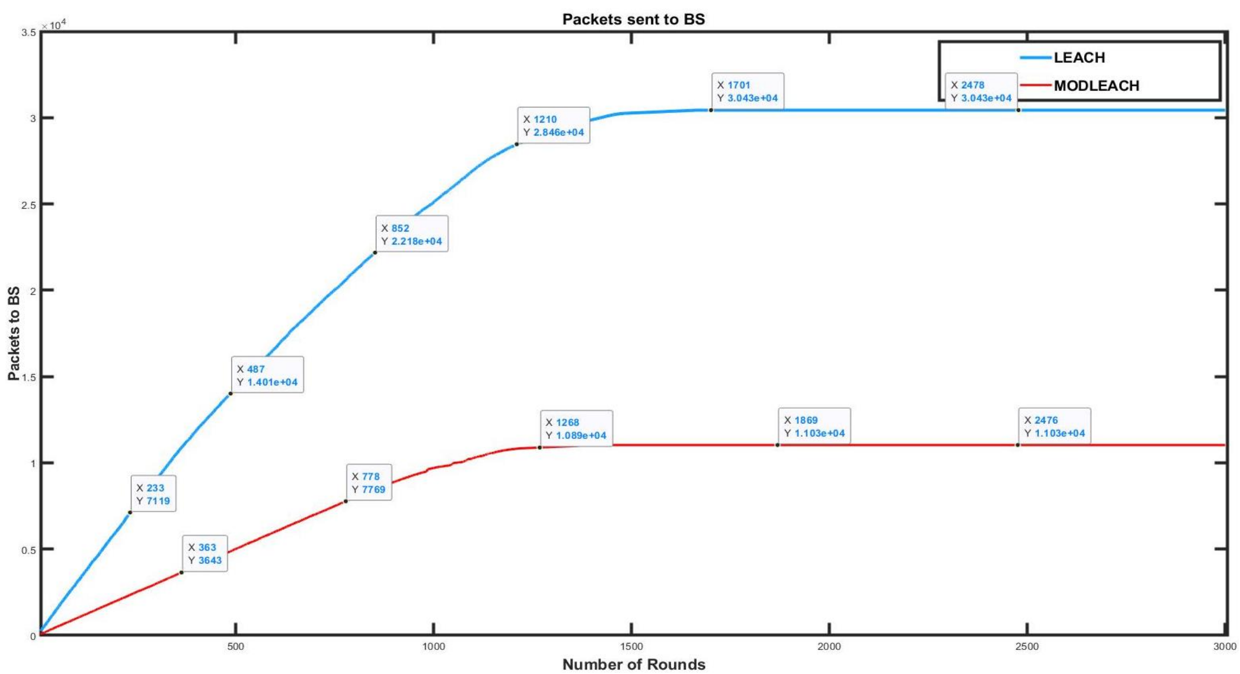

Figure 4-13: Packets to BS Comparison of LEACH and MODLEACH.

BS receives more packets from LEACH than MODLEACH by a factor of 2-to-3. It is somewhat surprising to see this result. A possible cause could be the soft and hard threshold settings for MODLEACH. Therefore, cluster formations are different than with LEACH, and distances are also different. A node's distance from the BS is enhanced with MODLEACH compared to LEACH, so more packets are transmitted with LEACH compared to two-threshold MODLEACH.

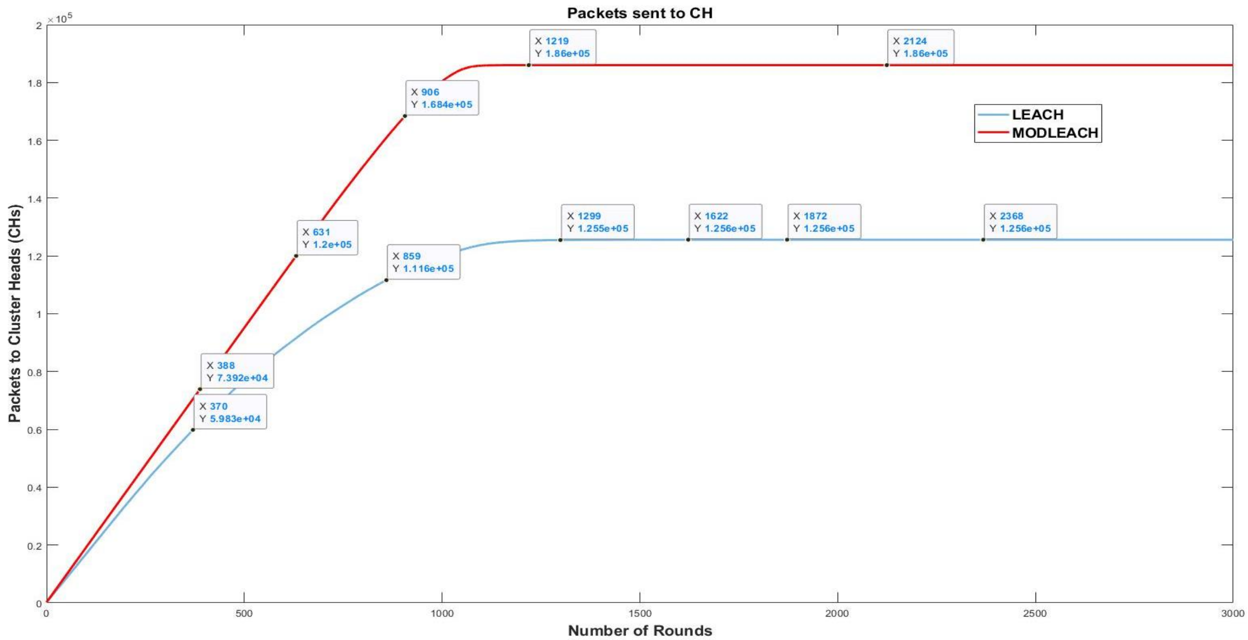

Figure 4-14: Packets to CH Comparison of LEACH and MODLEACH.

The results for packets transmitted to CH are more in MODLEACH as compared to LEACH by a factor of 2, a relatively consistent impact due to the enhanced performance of the MODLEACH algorithm by design compared to LEACH algorithm.

4.6 Progress Overview

Three distinct phases have been established for the project. The first is a literature review, a then the comparison of various algorithms in terms of results obtained from simulations, and, finally, an effort to create a brand-new algorithm from the evolutionary or genetic algorithm category.

4.7 Current progress

The modelling of WSN is performed in MATLAB, and sensors are randomly distributed in a given field of a specific area ranging from 100*100 and 300*300 m^2. The DEEC and its variations are contrasted in terms of essential factors like energy usage and network longevity. The outcomes of LEACH and MODLEACH are then compared, and MODLEACH's advancements are examined. The comparison can be expanded to other algorithms, such as TEEN, PEGASIS, DD, or well-known algorithms. However, the most important part is learning the trajectory for WSN deployment and studying how the different researchers approached this problem in the past.

4.8 Future Goals and Ambitions

The development of a novel algorithm and enhancements to the comparative analysis of various algorithms will be attempted in the following phase. The comparison of TEEN, PEGASIS and SEP is also underway and can be completed based on the successful completion of the code comparison, availability of time and resources. Modifications can be made to the properties of the sensor field, the parametric value of probabilities and their respective energy and power distribution at the TX and RX sides. These modifications may include combining comparisons in a single simulation and adjusting some parameters to obtain better energy consumption, network lifetime, and dead and alive nodes.

5. Results, Conclusions and Discussions

Due to the improved threshold mechanism, the MODLEACH exceeded the LEACH in performance. The operational/functional nodes against the rounds count for MODLEACH are very sharp sloped compared to LEACH, which is spread across the number of rounds and has a considerable slope value. In practically every area, including dead nodes, alive nodes, improved energy usage, and network longevity, DEEC outperforms its derivatives, such as DDEEC and EDEEC.

6. Future Recommendations

One area of improvement for EA-based routing protocols is scalability. As WSNs continue to grow in size and complexity, the ability of these protocols to handle large numbers of nodes and diverse network conditions will become increasingly important. Researchers may investigate new EA techniques or modifications to existing protocols that can improve scalability.

Another area of focus for future research is security. WSNs are often deployed in critical or sensitive environments; therefore, it's essential to ensure the security of data transmission and node communication security. Researchers may investigate new security mechanisms or modifications to existing protocols that can improve security.

Moreover, integrating Machine Learning techniques with routing protocols can be a promising field of research. Machine learning algorithms can learn from the network's characteristics and adjust the routing strategy accordingly, leading to more efficient and reliable routing in WSNs, Ding et al., (2021).

Finally, the use of mobile nodes and mobile sinks in WSNs can also be an exciting area of research. This can add more complexity and dynamic to the network and lead to new challenges in routing, J. Wang et al., (2020).

Overall, routing in multi-hop WSNs using evolutionary algorithms, LEACH variations, TEEN PEGSIS, and SEP will require developing new techniques and modifications to existing protocols that can improve scalability, security, and adaptation to dynamic and mobile contexts.

REFERENCES

Abidi, W., & Ezzedine, T. (2017). *New Approach for Selecting Cluster Head based on LEACH Protocol for Wireless Sensor Networks*. https://doi.org/10.5220/0006336101140120

Ahmed, S., Sandhu, M. M., Amjad, N., Haider, A., Akbar, M., Ahmad, A., Khan, Z., Qasim, U., & Javaid, N. (2013). iMOD LEACH: improved MODified LEACH Protocol for Wireless Sensor Networks. *Journal of Basic and Applied Scientific Research (JBASR) (ISSN 2090-4304)*, *3*, 25–32.

Ali, Q. I. (2012). Simulation Framework of Wireless Sensor Network (WSN) Using MATLAB/SIMULINK Software. In V. N. Katsikis (Ed.), *MATLAB* (p. Ch. 12). IntechOpen. https://doi.org/10.5772/46467

Ali QasimAlrubaye, I. (2022). Improvement of Cluster Head Selection in LEACH for Reducing Energy Consumption in Wireless Sensor Networks. *Journal*, *8*(1), 49–57.

Arumugam, G. S., & Ponnuchamy, T. (2015). EE-LEACH: development of energy-efficient LEACH Protocol for data gathering in WSN. *EURASIP Journal on Wireless Communications and Networking*, *2015*(1), 76. https://doi.org/10.1186/s13638-015-0306-5

Attiah, A., Chatterjee, M., & Zou, C. C. (2017). A Game Theoretic Approach for Energy-Efficient Clustering in Wireless Sensor Networks. *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 1–6. https://doi.org/10.1109/WCNC.2017.7925668

Bakr, B. A., & Lilien, L. (2011). LEACH-SM: A protocol for extending wireless sensor network lifetime by management of spare nodes. *2011 International Symposium of Modeling and Optimization of Mobile, Ad Hoc, and Wireless Networks*, 375. https://doi.org/10.1109/WIOPT.2011.5930046

Behera, T. M., Samal, U. C., & Mohapatra, S. K. (2018). Energy-efficient modified LEACH protocol for IoT application. *IET Wireless Sensor Systems*, *8*(5), 223–228. https://doi.org/https://doi.org/10.1049/iet-wss.2017.0099

Bhattacharyya, D., Kim, T., & Pal, S. (2010). A Comparative Study of Wireless Sensor Networks and Their Routing Protocols. *Sensors*, *10*(12), 10506–10523. https://doi.org/10.3390/s101210506

Cayirpunar, O., Kadioglu-Urtis, E., & Tavli, B. (2015). Optimal Base Station Mobility Patterns for Wireless Sensor Network Lifetime Maximization. *IEEE Sensors Journal*, *15*(11), 6592–6603. https://doi.org/10.1109/JSEN.2015.2463679

Chaurasiya, S. K., Biswas, A., & Banerjee, R. (2021). Metaheuristic Multilevel Heterogeneous Clustering Technique for Heterogeneous Wireless Sensor Networks. *2021 10th International Conference on Internet of Everything, Microwave Engineering, Communication and Networks (IEMECON)*, 1–6. https://doi.org/10.1109/IEMECON53809.2021.9689098

Chen, C., Wang, L.-C., & Yu, C.-M. (2022). D2CRP: A Novel Distributed 2-Hop Cluster Routing Protocol for Wireless Sensor Networks. *IEEE Internet of Things Journal*, *9*(20), 19575–19588. https://doi.org/10.1109/JIOT.2022.3148106

Ding, Q., Zhu, R., Liu, H., & Ma, M. (2021). An Overview of Machine Learning-Based Energy-Efficient Routing Algorithms in Wireless Sensor Networks. *Electronics*, *10*(13). https://doi.org/10.3390/electronics10131539

Fanian, F., Kuchaki Rafsanjani, M., & Bardsiri, V. (2016). A Survey of Advanced LEACH-based Protocols. *International Journal of Energy, Information and Communications*, *7*, 1–16. https://doi.org/10.14257/ijeic.2016.7.1.01

Geeksforgeeks. (2021, June 3). *Wireless Sensor Network (WSN)*. https://www.geeksforgeeks.org/wireless-sensor-network-wsn/

He, L.-M. (2009). A Novel Real-Time Routing Protocol for Wireless Sensor Networks. *2009 10th ACIS International Conference on Software Engineering, Artificial Intelligences, Networking and Parallel/Distributed Computing*, 411–416. https://doi.org/10.1109/SNPD.2009.110

Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, 10 pp. vol.2-. https://doi.org/10.1109/HICSS.2000.926982

Huruială, P.-C., Urzică, A., & Gheorghe, L. (2010). Hierarchical routing protocol based on evolutionary algorithms for Wireless Sensor Networks. *9th RoEduNet IEEE International Conference*, 387–392.

Kandpal, R., Singh, R., & Mandoria, H. L. (2015). *Analysis on Enhancements in LEACH Protocol for WSN*.

Khan, A. R., Rakesh, N., Bansal, A., & Chaudhary, D. K. (2015a). Comparative study of WSN Protocols (LEACH, PEGASIS and TEEN). *2015 Third International Conference on Image Information Processing (ICIIP)*, 422–427. https://doi.org/10.1109/ICIIP.2015.7414810

Khan, A. R., Rakesh, N., Bansal, A., & Chaudhary, D. K. (2015b). Comparative study of WSN Protocols (LEACH, PEGASIS and TEEN). *2015 Third International Conference on Image Information Processing (ICIIP)*, 422–427. https://doi.org/10.1109/ICIIP.2015.7414810

Kiran Kumar Panigrahi. (2022, July 28). *Difference between Star and Mesh Topology*. www.tutorialspoint.com/difference-between-star-and-mesh-topology

Koltsidas, G., & Pavlidou, F.-N. (2011). A game theoretical approach to clustering of ad-hoc and sensor networks. *Telecommunication Systems*, *47*(1), 81–93. https://doi.org/10.1007/s11235-010-9303-5

Kulkarni, N., Prasad, N., & Prasad, R. (2013). MOHRA: Multi Objective Hybrid Routing Algorithm for Wireless Sensor Network. *Wireless VITAE 2013*, 1–6. https://doi.org/10.1109/VITAE.2013.6617056

Lakshmi, P. S., Jibukumar, M. G., & Neenu, V. S. (2018). Network lifetime enhancement of multi-hop wireless sensor network by RF energy harvesting. *2018 International Conference on Information Networking (ICOIN)*, 738–743. https://doi.org/10.1109/ICOIN.2018.8343216

Liu, Q., & Liu, M. (2017). Energy-efficient clustering algorithm based on game theory for wireless sensor networks. *International Journal of Distributed Sensor Networks*, *13*(11), 1550147717743701. https://doi.org/10.1177/1550147717743701

Liu, X., Liu, Y., & Bai, T. (2008). Energy-Efficient Real-Time Routing in Wireless Sensor Networks. *2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application*, *2*, 1009–1013. https://doi.org/10.1109/PACIIA.2008.295

Lu, Y., Liu, X., & Li, M. (2017). *Study on Energy-Saving Routing Algorithm Based on Wireless Sensor Network*.

Mahmood, D., Javaid, N., Mahmood, S., Qureshi, S., Memon, A., & Zaman, T. (2013). *MODLEACH: A Variant of LEACH for WSNs*. https://doi.org/10.1109/BWCCA.2013.34

Manjeshwar, A., & Agrawal, D. P. (2001). TEEN: a routing protocol for enhanced efficiency in wireless sensor networks. *Proceedings 15th International Parallel and Distributed Processing Symposium. IPDPS 2001*, 2009–2015. https://doi.org/10.1109/IPDPS.2001.925197

Maurya, P., & Kaur, A. (2016). A Survey on Descendants of LEACH Protocol. *International Journal of Information Engineering and Electronic Business*, *8*, 46–58. https://doi.org/10.5815/ijieeb.2016.02.06

Mittal, N., Singh, U., Salgotra, R., & Sohi, B. (2018). A boolean spider monkey optimization based energy efficient clustering approach for WSNs. *Wireless Networks*, *24*. https://doi.org/10.1007/s11276-017-1459-4

Nam, D. (2020). Comparison Studies of Hierarchical Cluster-Based Routing Protocols in Wireless Sensor Networks. *CATA*, *69*, 334–344.

Nayak, P., & Vathasavai, B. (2017). Energy Efficient Clustering Algorithm for Multi-Hop Wireless Sensor Network Using Type-2 Fuzzy Logic. *IEEE Sensors Journal*, *17*(14), 4492–4499. https://doi.org/10.1109/JSEN.2017.2711432

Omari, M., & Laroui, S. (2015). Simulation, comparison and analysis of Wireless Sensor Networks protocols: LEACH, LEACH-C, LEACH-1R, and HEED. *2015 4th International Conference on Electrical Engineering (ICEE)*, 1–5. https://doi.org/10.1109/INTEE.2015.7416826

Qureshi, T. N., Javaid, N., Malik, M., Qasim, U., & Khan, Z. A. (2012). On Performance Evaluation of Variants of DEEC in WSNs. *2012 Seventh International Conference on Broadband, Wireless Computing, Communication and Applications*, 162–169. https://doi.org/10.1109/BWCCA.2012.35

Rahmadhani, M. A., Yovita, L. v, & Mayasari, R. (2018). Energy Consumption and Packet Loss Analysis of LEACH Routing Protocol on WSN Over DTN. *2018 4th International Conference on Wireless and Telematics (ICWT)*, 1–5. https://doi.org/10.1109/ICWT.2018.8527827

Redjimi, K., Mehdi, B., & Mohamed, R. (2022). *DEEC and EDEEC Routing Protocols for Heterogeneous Wireless Sensor Networks: A Brief Comparative Study* (pp. 117–125). https://doi.org/10.1007/978-3-030-98531-8_12

Saini, P., & Sharma, A. K. (2010). E-DEEC- Enhanced Distributed Energy Efficient Clustering scheme for heterogeneous WSN. *2010 First International Conference On Parallel, Distributed and Grid Computing (PDGC 2010)*, 205–210. https://doi.org/10.1109/PDGC.2010.5679898

Saleem, M. M., & Alabady, S. A. (2022). Improvement of the WMSNs lifetime using multi-hop clustering routing protocol. *Wireless Networks*, *28*(7), 3173–3183. https://doi.org/10.1007/s11276-022-03036-3

Samaras, N. S., & Triantari, F. S. (2016). On Direct Diffusion Routing for Wireless Sensor Networks. *2016 Advances in Wireless and Optical Communications (RTUWO)*, 89–94. https://doi.org/10.1109/RTUWO.2016.7821862

Sefuba, M., & Walingo, T. (2018). Energy-efficient medium access control and routing protocol for multihop wireless sensor networks. *IET Wireless Sensor Systems*, *8*(3), 99–108.

Shagari, N. M., Idris, M. Y. I., Salleh, R. B., Ahmedy, I., Murtaza, G., & Shehadeh, H. A. (2020). Heterogeneous Energy and Traffic Aware Sleep-Awake Cluster-Based Routing Protocol for Wireless Sensor Network. *IEEE Access*, *8*, 12232–12252. https://doi.org/10.1109/ACCESS.2020.2965206

Sharma, R., Vashisht, V., & Singh, U. (2020). Modelling and simulation frameworks for wireless sensor networks: a comparative study. *IET Wirel. Sens. Syst.*, *10*, 181–197.

Singh, A., Sharma, S., & Singh, J. (2021). Nature-inspired algorithms for Wireless Sensor Networks: A comprehensive survey. *Computer Science Review*, *39*, 100342. https://doi.org/https://doi.org/10.1016/j.cosrev.2020.100342

Song, P. (2013). Multi-Hop Reliable Transmission Algorithm Based on Optimal Selection and Forwarding in Wireless Sensor Network. *2013 International Conference on Computational and Information Sciences*, 1429–1432. https://doi.org/10.1109/ICCIS.2013.377

Taylor, K. ,. (2022, November 11). *What is the Importance of Predictive Maintenance in Industry 4.0?*

Technologify. (2021). *Wireless Sensor Networks – Definition, Applications, Design Issue, and More*. https://www.technologyify.com/wireless-sensor-networks/

Uday B. Mavdiya, & Limbad, N. (2015). A Literature Survey of Energy-Efficient Clustering in WSN Using Different Variant of LEACH Routing Protocol . *International Journal for Scientific Research & Development*, *3*(01), 530–532.

Wang, G., Zhu, H., Dai, H., Wu, L., & Xiong, B. (2009). The Clustering Algorithm of Wireless Sensor Networks Based on Multi-hop between Clusters. *2009 WRI World Congress on Computer Science and Information Engineering*, *3*, 177–181. https://doi.org/10.1109/CSIE.2009.593

Wang, J., Gao, Y., Zhou, C., Sherratt, S., & Wang, L. (2020). Optimal coverage multi-path scheduling scheme with multiple mobile sinks for WSNs. *Computers, Materials & Continua*, *62*(2), 695–711.

Xie, D., Sun, Q., Zhou, Q., Qiu, Y., & Yuan, X. (2013). An Efficient Clustering Protocol for Wireless Sensor Networks Based on Localized Game Theoretical Approach. *International Journal of Distributed Sensor Networks*, *2013*. https://doi.org/10.1155/2013/476313

Xin Qu. (2012). *Energy Efficient Wireless sensor Networks with Modified LEACH Algorithm* . Blekinge Institute of Technology, Sweden.

Xue, Q., & Ren, X. (2012). Research of routing protocols simulation for wireless sensor networks based on OMNeT++. *2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, 79–82. https://doi.org/10.1109/ICQR2MSE.2012.6246191

Yadav A., & Kumar S. (2016). An Enhanced Distributed Energy-Efficient Clustering (DEEC) Protocol for Wireless Sensor Networks . *International Journal of Future Generation Communication and Networking*, *9*(11), 49–58.

Yagouta, A. ben, Gantassi, R., & Gouissem, B. ben. (2017). Compromises between energy consumption and quality of service metrics in wireless sensor networks with mobile sink and cluster based routing protocols. *2017 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC)*, 60–66. https://doi.org/10.1109/IINTEC.2017.8325914

Yagouta, A. ben, Jabberi, M., & Gouissem, B. ben. (2017). Impact of Sink Mobility on Quality of Service Performance and Energy Consumption in Wireless Sensor Network with Cluster Based Routing Protocols. *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*, 1125–1132. https://doi.org/10.1109/AICCSA.2017.198

Yang, T., Barolli, L., Mino, G., Xhafa, F., & Durresi, A. (2010). Impact of Mobile Event Movement on the Performance of Wireless Sensor Networks. *2010 Fifth International Conference on Systems and Networks Communications*, 88–93. https://doi.org/10.1109/ICSNC.2010.19

Yao, Y.-D., Li, X., Cui, Y.-P., Deng, L., & Wang, C. (2022). Game Theory and Coverage Optimization Based Multihop Routing Protocol for Network Lifetime in Wireless Sensor Networks. *IEEE Sensors Journal*, *22*(13), 13739–13752. https://doi.org/10.1109/JSEN.2022.3178441

Yi, X., Deng, L., & Liu, Y. (2010). Multi-hop Clustering Algorithm for Wireless Sensor Networks. *2010 International Conference on Measuring Technology and Mechatronics Automation*, *2*, 728–731. https://doi.org/10.1109/ICMTMA.2010.319

Zhang, J., & Yan, R. (2019). Centralized Energy-Efficient Clustering Routing Protocol for Mobile Nodes in Wireless Sensor Networks. *IEEE Communications Letters*, *23*(7), 1215–1218. https://doi.org/10.1109/LCOMM.2019.2917193

Zhao, L., Qu, S., & Yi, Y. (2018). A modified cluster-head selection algorithm in wireless sensor networks based on LEACH. *EURASIP Journal on Wireless Communications and Networking*, *2018*(1), 287. https://doi.org/10.1186/s13638-018-1299-7

Zhihui, H. (2015). Research on WSN Routing Algorithm Based on Energy Efficiency. *2015 Sixth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA)*, 696–699. https://doi.org/10.1109/ISDEA.2015.178

11/27/2022

Deep Learning Techniques For Predictive Maintenance In Industry 4.0.

User

FOURTH EDITED VERSION

ABSTRACT

Cognitive Radio (CR) was introduced in 1999, with potential applications such as Software Defined Radios (SDRs) and reconfigurable radios over wireless networks. Communication between intelligent radio devices and associated network entities is described in it. As they learn from their experience, they can tune or change their operating parameters per the network demands. The wireless community has been enthused by such a concept that attempts to mimic human cognition and reasoning, which has sparked many research and standardization activities.

Federal Communication Commission (FCC) efforts have focused on allocating and starting the spectrum more efficiently using intelligent schemes. Between 2002 and 2010, the FCC issued rules and regulations allowing unlicensed devices to exploit TV white space opportunistically. This initiative has sparked an avalanche of spectrum measurement campaigns worldwide, aiming to demonstrate spectrum underutilization.

Since the focus is on the experimental study through simulations, the study started with creating multiple topologies. The starting research is done with four topologies, and each topology is tested with different types of transmitters and receivers, each topology having its own data types, hence different data types for each topology, each topology having its own properties and protocols.

The selection of simulators is also a crucial step towards successfully implementing the research study. Network simulators play a pivotal role in the research process. Research has been conducted to evaluate the performance of different routing protocols [2] in other simulators [3] with varying network parameters. The nodes in a WSN can be simulated with MATLAB scripts, and their relevant energy consumption, network lifetime, node coverage, and clustering-based nodes (CHs and Sinks) can be plotted. However, MATLAB is not the first choice for Cognitive Radio Network simulations. A more commonly adopted preferred option is NS-2 and NS-3. Some other simulators include GloMoSim/QualNet, OMNeT++, TOSSIM, OPNET Modeler Wireless Suite, MiXiM, Castalia, INET framework, NesCT, Avrora, NS-2, J-SIM, ATEMY, Emstar, SENS, SENSE, and SHAWN., Ali, (2012)

List of Tables

List of Figures

List of Acronyms and abbreviations

CRN	Cognitive Radio Network
SDR	Software Defined Radio
CR	Cognitive Radio
SCADA	Supervisory control and data acquisition
PrM	Proactive Maintenance
PM	Preventive Maintenance
PdM	Predictive maintenance
CMMS	Computerized maintenance management system
IIoT	Industrial Internet-of-things
AI	Artificial Intelligence
CM	Corrective Maintenance
TBM	time-based maintenance
CBM	Condition-Based Maintenance
ML	Machine Learning
LSTM	Long Short-Term Memory
RNN	Recurrent Neural Network
CNN	Convolutional Neural Network

RF	Random Forest
ReLu	Rectified Linear Unit
MPM	Maintenance Policy Management
FMEA	Failure Modes and Effects Analysis
FTA	Fault Tree Analysis
DNN	Deep Neural Network
ANN	Artificial Neural Network

Table of Contents

1. Introduction

Cognitive radio networks (CRNs) divide their users into two classes: the main users are referred to as primary users (PUs), and to coexist with them; cognitive radios (CRs) also provide secondary users (SUs) that share the same spectrum resources with primary users without causing major interference. Innovative/smart policies enable SUs to intelligently utilize spectrum resources in CRNs by automatically sensing licensed channels authorized to PUs. Based on smart opportunistic access, CR technology integrates traditional wireless networks with new technologies. CRs rely on the composition of wireless networking protocols layer by layer to form a robust communicational network. On the other hand, they extend frequency resource availability by using opportunistic spectrum sharing (OSS).

2. Literature Review

Keywords:

Internet of Things (IoT), Big Data, Industry 4.0, cyber-physical systems (CPS), Augmented Reality, Artificial Intelligence, Machin learning

This chapter presents a systematic literature review of Deep Learning applications in operational decision-making using PdMin Industry 4.0.

A recent study from Deloitte resulted in information that a PdMstrategy can result in a reduction of time wasted on maintenance planning by 20-50%, enhance the efficiency in utilization of the equipment by 10-20% and decrease the overall maintenance budget by 5-10% (Spronk, 2022, p. 2). There are also results from Forbes studies that a plant's downtime can cost 45-50 billion € (44.8-50 billion $) a year. The same research also indicates that shutdowns can waste 1-10% of peak production time, whether planned or not (Sundeep R., 2022, p. 3).

2.1 Maintenance Philosophies

According to (Giuliano Liguori, 2022, pp. 1–3), there are two kinds of maintenance, corrective, also termed as reactive and preventive.

Corrective maintenance is restoring equipment, system

, or machine to its primary and genuine working and original condition.

The object, apparatus, or method is used until a failure happens. This incurs unreliability as the time of failure of equipment is unpredictable and unknown in advance, which causes system downtime (economic loss) and safety problems (temporal loss).

Preventive maintenance is maintaining equipment

through specific processes to avoid failures from happening.

Corrective maintenance is costly and more prone to economic and equipment loss compared to preventive.

With the emergence of Artificial intelligence, Machine learning, Data Science and analytics, Internet-of-things, digital age, another form of maintenance is established as the most popular form at the highest decision-making level.

PdMis a branch of computer science/engineering that uses the massive power of data generated from billions of devices, applying filtering and

data cleaning techniques on data, and extracting results from

that data to drive future-based decision-making to

predict in advance the failures that are going to occur in systems.

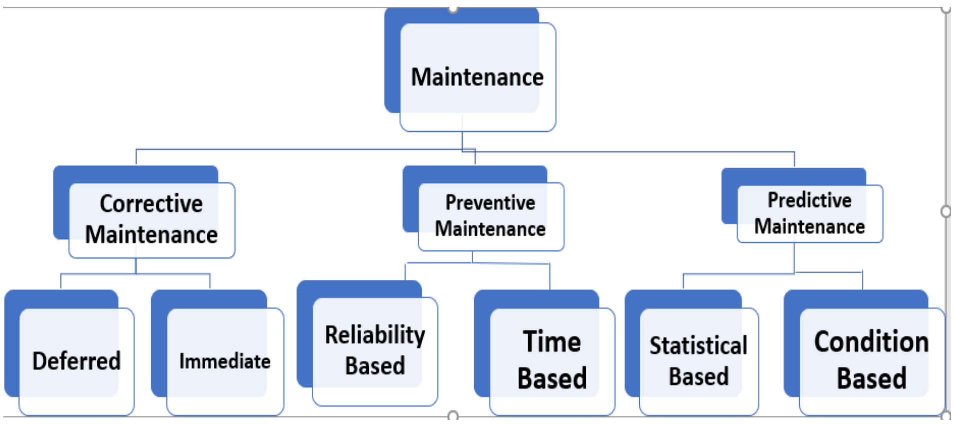

Figure 2-1: Maintenance Philosophies

2.1.1 Condition-based maintenance

Condition-based maintenance (CBM) systems find many applications in niches such as automotive, manufacturing, defence, aviation, oil plants, and power distribution systems. It is a good idea to read about the optimization of CBM-based maintenance strategy and study its implications in society and benefits such as cost reduction in automotive and automatic fault diagnosis and prognosis using alerts for fault detection (Ashok Prajapati, 2012, pp. 384-400).

In CBM, the management and repair decisions are made based on the current condition and state of the equipment or machine. This state is accessed by collecting diagnostics and prognosis data on equipment using sensors and multiple methods such as vibration analysis, infrared thermography current, oil analysis, etc. Different maintenance terminologies and strategies are reliability-centred maintenance (RCM), total productive maintenance (TPM), condition-based maintenance (CBM), computerized maintenance management systems (CMMS), auditing systems, etc. (Ellis & Byron, 2008, pp. 1–4). The equipment's current condition determines the exact time for performing maintenance operations on the assets or subcomponents. Hence, modelling the asset's condition becomes a crucial and challenging problem; an excellent attempt has been made to model the CBM model for water pumps using vibrations of bearings at six varying positions (W. Wang et al., 2000).

Industry 4.0, smart manufacturing, and digital factory are the concepts that have been gaining attention recently. There are multiple stages in the prediction of the failure of the equipment, one of which is *Diagnosis* which is the detection of the failure by using heterogenous data to detect abnormal operations or processes of the equipment; the next stage can be prognosis, which is predicting future failure times of the equipment, once the detection and failure times information is known, the decision-making stage can be gone ahead (proactive decision making), for making decisions in advance to prevent significant economic and equipment losses. The paper (Bousdekis et al., 2019, pp. 1–3) presents a systematic literature review with the following points, maintenance planning, scheduling, decision-making based on reliability, and time-based degradation of equipment. Optimization techniques, such as joint optimization, multi-component, multistate system optimization, and estimation of risk and maintenance cost-based optimization.

The fourth industrial revolution has given birth to many new technologies, one of which is predictive maintenance, a crucial element for sustainable manufacturing and industrial revolution, the concept is equivalent to the digital version of machine maintenance. As sensing technologies revolutionize, the data generated from production plants and processes has grown exponentially. Machine learning and deep learning are both sub-categories of Artificial intelligence. Machine learning algorithms are categorized into three categories which are supervised learning, unsupervised learning, and reinforcement learning (Julianna D., 2021, p. 2). Supervised learning takes advantage of the structure of data (or uses labeled data) to create classes (classification) and apply regression. Unsupervised learning uses unlabeled data, and clusters the data based on similarity in the data of neighbors. Supervised learning algorithms are from classification and regression such as artificial neural networks, decision trees, ensemble methods, Support Vector Machines (SVM), Naïve Bayes, nearest neighbour, and discriminant analysis, while unsupervised algorithms are hidden Markov model and neural network-based (Z. Cinar et al., 2020, p. 5). The main machine learning algorithms for industry 4.0 are data acquisition, classification, clustering, etc. PHM (prognostics and health management), as it is called in I4.0 (industry 4.0), is crucial for a sustainable industry and a reliable source for equipment health monitoring. PdMcan achieve no failures, near-zero unseeable dangers, no accidents, and 0% pollution. PdM(PdM) is categorized into three branches, 1) prognosis based on a model, 2) prognosis based on knowledge, and 3) prognosis based on data-driven techniques. IWSNs (industrial wireless sensor networks) have become standard for data acquisition in cyber-physical systems (CPS) (W. Zhang et al., 2019, p. 1).

An intelligent PdM methodology was demonstrated by (Abidi et al., 2022, pp. 1–6). This model comprises six steps: cleaning of data, normalization of data, optimal feature selection, accurate network decision-making, and finally, prediction stage. The target is to forecast the next failure and take preventive measures before the subsequent failure happens. Two practical datasets were used, and a five-step method was adopted to reach prediction, data cleaning, normalization, optimal feature selection (by combining Jaya algorithm and Sea Lion Optimization (SLnO), prediction network decision-making, and prediction. SVM is used to provide decisions on the prediction network. The forecast is eventually performed using RNN (recurrent neural network), and weight optimization is performed using combined J-SLnO. The authors compared the 8 papers using different machine learning methods such as Random Forest, SVM, LSTM (long short-term memory), NN (neural network), elbow method, and RNN (recurrent neural network).

The types of predictions that an intelligent performing PdMsystem can make are:

i. Remaining useful life (RUL): This prediction metric helps determine the exact time when the equipment is expected to fail; hence, in advance, maintenance can be scheduled. This can be done with the help of regression techniques.
ii. Flagging irregular behaviour: This can be done with the help of anomaly detection techniques from machine learning, by using time series analysis.
iii. Two-step procedure for failure prediction: 1) diagnosis and 2) suggestions of reduction or maintenance action after the failure occurs. This can be implemented using classification techniques from ML/DL.

Machine learning and deep learning enable us to:

i. Design creative models that can increase the operational efficiency and uptime of plant or equipment and enhance the asset's lifetime.
ii. ii. The cycle maintenance procedures can be optimised using historical and current data.
iii. Reduction or elimination of downtime can help save customers, money, and probably precious lives.

The history data is a good indication of the prediction of Remaining useful life (RUL) or failure diagnosis (Gonfalonieri, 2019, pp. 1–3). Some examples of enterprise software for PdMare **CMMS** Computerized Maintenance Management System (or Software), also termed Enterprise Asset Management (EAM) software.

Based on scenarios, the maintenance has also been categorized by researchers in the following categories: proactive Maintenance (PrM), preventive maintenance (PM), and PdM (Predictive maintenance). Six sigma is another system invented by Motorola in the late 1970s and early 1980s for fault detection analysis. When General Electric adopted it, it became well-known. The acceptable degree of perfection for six sigma is 3.4 errors for every billion opportunities.The question arises if the industry in consideration can afford that level of excellence or is overkill except for mission-critical components, systems, and industries. The four dimensions of the PM (preventive maintenance) system and six patterns of failures are detailed by(J. Levitt, 2011, pp. 31–72). The same book in chapter 11 links condition-based maintenance to predictive maintenance. Several primary diagnostic tools used for PdMin steel manufacturing by Gary works are Vibration Analysis, Thermography, fluid analysis, visual inspection, Operational-dynamics analysis, electrical monitoring, and failure analysis (J. Levitt, 2011, pp. 141–147).

The school of thought on PdMvaries as per industry. The equipment can be classified as:

Critical: This class of equipment is costly, has high repair costs, and ample time to repair this type of equipment (e.g., high-speed turbomachinery), and failure of such assets can hamper plant safety and stop production.

Essential: In this category of assets, defects can affect plant safety; the equipment requires intermediate expenditure as repair costs, low-medium level expertise, and time to repair.

General purpose: Not critical equipment as stoppage or failure of such pieces of equipment does not affect safety.

Why PdMworks because a conservative approach to maintenance is regular check-ups and maintenance. This is costly as the tires and engine are changed before they complete their life cycle, or over-deterioration can damage the equipment, resulting in a more significant loss. PdM is calculated and resource-efficient, as it learns from data and predicts the equipment failure time. The two extremes of the conservative approach are avoided. The perfect dataset required for PdMis time-series data. For a good PdM model, the following considerations are important.

i. The failures needed to predict must be understood by the model designer. Which failures are required to be anticipated?
ii. Is the failure happening over time or is it abruptly occurring? Each failure will have its associated processes and parameters. Each of these parameters will have to be monitored using specific sensors.
iii. Every piece of equipment normally has a specific working period; the data thus needs to be collected over a long period for proper failure diagnosis (Mitul M., 2020, pp. 2–3).

According to research from Market Research Future, the PdMmarket will increase by at least 25% CAGR (Compound Annual Growth Rate) and reach the 23 million USD milestone by 2025. PdM is thought to be the most cutting-edge use of Internet of Things technologies. The most popular ML technique is anomaly detection, which over time, learns a machine's typical behaviour and raises the alarm in the event of aberrant behaviour. Some key concepts must be comprehended to apply ML solutions successfully.

- History of errors/problems of machine
- Machine Maintenance & repair history
- Equipment proper operating conditions
- Machines metadata

IBM service for PdM is called IBM PdMand Quality, and Azure and AWS provide similar services. Traditional managers implemented PdM with the help of SACDA, but it demanded manually coded thresholds and alert systems that don’t offer dynamically changing data of the equipment processes (Chuprina R, 2021, p. 4).

2.2 Industry 4.0

The first industrial revolution was mechanized production, the progress of manual output, such as the steam engine and hydropower, and the emergence of the textile industry. As a result of the second revolution, railroad networks became large-scale. Digital electronics and the internet became dominant during the third industrial revolution. The widespread adoption of the Internet-of-things, artificial intelligence, data analytics, intelligent decision-making, automation, cyber-physical systems, smart homes, cities, and interconnected roads characterizes the fourth industrial revolution.

The part of the internet-of-things, when applied in an industrial scenario, is called industrial internet-of-things, when combined with artificial intelligence, big data analytics, and edge/cloud computing-based decision-making, constitute predictive maintenance. These technologies are under the umbrella of industry 4.0, making PdM an integral part of industry 4.0 (vroc.ai team, 2022, pp. 1–2).

Links must be established between the IoT, IIoT, machine learning, and deep learning applications in an industrial context, PdM, and making industry 4.0 a reality to make PdM easier to understand.

Intelligent PdMis enabled by two of the biggest revolutions in technology, namely LoTand Artificial Intelligence. PdM is a branch of industry 4.0, which monitors the change in equipment parameters or process through a defined set of activities and alerts the system to take maintenance decisions in real-time, maximizing the lifetime of the equipment. A typical process in industry 4.0 for intelligent PdMworks by collecting data generated from CPS and IoT and processing through different stages (using deep learning to find patterns or abnormal behaviours) and alerting about a possible future fault through a data mining system for maintenance decision-making. Hence, a systematic analysis of computational intelligence (Artificial Neural Networks (ANN), Fuzzy Logic Systems (FLS) and Evolutionary Computing (Genetic Algorithms)), ANN, DNN, RNN, CNN, and deep learning is of utmost importance (K. Wang & Wang, 2018, pp. 2–7).

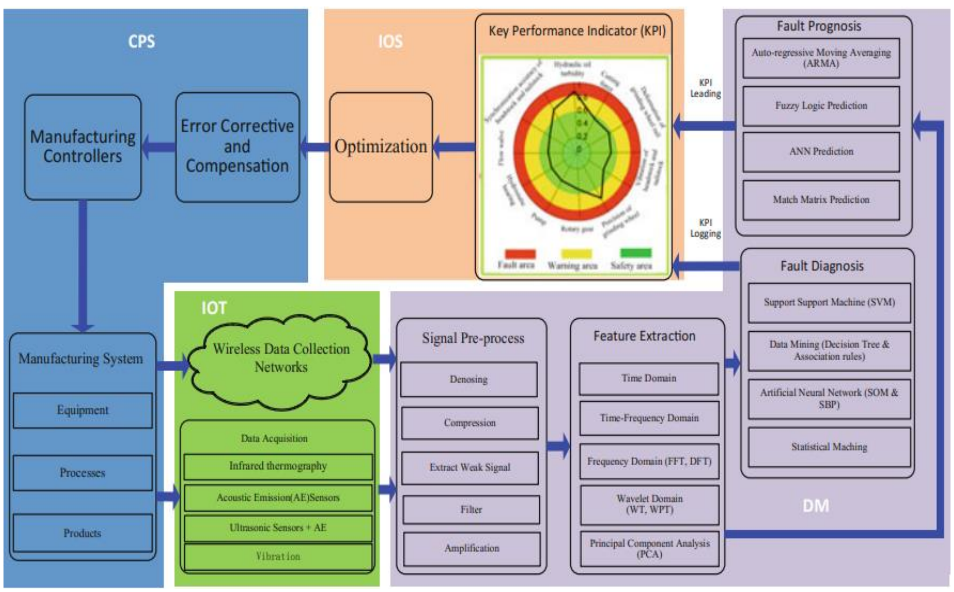

Figure 2-2: An intelligent PdM(IPdM) system in industry 4.0*(K. Wang & Wang, 2018)*.

2.3 Industrial Internet-of-Things

Kevin Ashton conceived the IoTconcept in 1999. the IoT was for the commercial sector, while IIoT was built for industrial applications. IIoT includes the IoT where IoT devices in an industry interact with existing set-ups to increase efficiency and productivity in manufacturing sectors such as cement, gas, oil, energy, etc. Industry 4.0 provides the leap necessary for the growth of the industrial Internet of Things. According to the first definition of IloT by *General Electric*: "IloT comprises of two major areas; the connection of components to local processing and the internet; and further to other industrial networks that can be independently used to generate value or revenue. Industrial Internet-of-Things, (IIoT) is a critical industrial IoT infrastructure and aids in securing networks in the industry for ultra-critical applications. An introduction to internet technologies in the distributed control systems (DCS), & Industrial control systems (ICS), composed of operational technology (OT) and manufacturing sectors and plants, are presented in (Supreeth, A, 2020, pp. 2-3). The following are IoT applications for the energy sector.

1. Analysis of thousands of sensors and data points for efficiency gains.
2. Remote solar panel monitoring and maintenance.
3. Hazardous analysis of nuclear facilities
4. Smart electric, gas, and water meters in a citywide deployment to monitor usage and demand.
5. Time-of-use tariffs.
6. Real-time blade adjustments as a function of weather on remote wind turbines.

Now the internet-of-things, applied in the industry became IIoT, and the next thing is to implement deep learning in an industrial environment to efficiently extract the results from sensor data. Another literature review outlines the latest deep learning that is being researched in today's literature regarding PdMby taking into consideration an important concept, fast PdM, trying to optimize the latency in sensor data transfer to edge, cloud, or base station and deep learning inference model and make a fast real-time system as a whole. The latency for real-time can be in microseconds, or milliseconds based on the application if it is critical. 5G applications require an ultra-reliable, low latency system, with less than 10 ms latency applications. This is where two main challenges are revealed in deep learning for predictive maintenance, the processing speed as most deep learning systems require GPU, resulting in high power consumption, and highly capable processing equipment and high processing time, resulting in not-so-fast or real-time communications. IoT, IIoT, or sensors are limited in processing power, battery capability, and fast real-time (millisecond) latency time for data transfer (Rieger et al., 2019, pp. 3–8).

The main applications for PdMare in mechanical and electrical equipment monitoring. This is why electromechanical sensors are important. Many types of sensors including, industrial electromechanical sensors are surveyed in detail and some potential future applications outside of the industrial context are researched in Table 2.1.

Table 2-1:Sensor types, analysis, algorithms and applications (Namuduri et al., 2020).

Sensor type	**Type of analysis**	**Algorithm**	**Applications**
Electromechanical sensors	Prediction	Deep learning	Predictive Maintenance
Environmental sensors	Anomaly detection	Neural Networks	Smart Manufacturing
Physical sensors	Prognostics	CNN	Industry 4.0
Acoustic sensors	Forecasting	RNN	Industrial IoT
Electrical sensors	Remaining useful life	LSTM	Health Monitoring
Image sensors	Trend Analysis	Auto Encoders	

The next step is data collection from sensors, as this data comes from thousands of sensors, it is in a multi-variant time series format. After this, deep learning architecture needs to be chosen. This is followed by steps such as pre-processing of data (denoising, dimensionality reduction), and anomalies are detected, which indicate the faults in the sensor values. A detailed review of anomaly detection can be found in (Domingues et al., 2018, pp. 1–3). Next, prognostics, forecasting, and remaining usable life (RUL) are carried out, which, respectively, imply anticipating future issues, forecasting future sensor value, and estimating the time in the future when sensor value will drift to a value comparable to an equipment malfunction or failure.(Namuduri et al., 2020, pp. 2–8).

2.4 Summary of Literature Review

Industry 4.0 research and PdM started with the 4th revolution after the internet, which was the internet-of-things, which promised to connect every aspect of life and thing to the internet. With this, the industry pursued the IoTfor its applications and implemented IIoT. Next followed is artificial intelligence, which asked for implementations of deep learning in the industry for applications such as predictive maintenance. This resulted in constraints and challenges such as low power of sensors, low processing power such as GPU for remote sensors, and high latency for transmission of data, from sensors to deep learning devices and back to the industrial control centre for results inference and maintenance decision making. Deep learning is systematically studied in the next chapter, and algorithms applicable to PdM are researched. Performance is analyzed so that a chain architecture-type understanding of the problem at hand can be achieved.

3. Deep learning in Predictive Maintenance

In this chapter, the machine learning model training procedure from the start to the end of the model training chain is systematically presented. The algorithms from deep learning and machine learning that are suitable for PdM applications in the context of industry 4.0 are reviewed.

3.1 Literature review on deep learning for predictive maintenance

Studying Prognostic Health and Management (PHM) literature, PdM is divided into two steps, 1) prognostics (prediction of the parameters such as RUL, failure time, etc.) and decision-making based on prognostic measurements. Taking this approach forward, PdM is classified into model-based and data-driven-based PdM frameworks. Model-based approaches need to know the equipment degradation process and model it correctly. Markov processes can be used in such applications. This approach does not work in a variable operation process, as the system parameters are changing and cannot be modelled accurately in advance. Hence data-based methods are preferred. The techniques used in the study (Nguyen & Medjaher, 2019, pp. 1–5) are LSTM, feature engineering, Convolutional Bi-directional Long Short-Term Memory networks (CBLSTM), restricted Boltzmann machine to extract the sensor parameters from devices, or appliances, or machines and extract RUL (remaining useful life) using feature engineering techniques.

Hybrid approaches combine the properties of data-based and physical (model) based methods. OSA-CBM is an open system architecture for condition-based monitoring and advises a six-step procedure for predictive maintenance implementation (OSA-CBM, 2022, pp. 1–2).

Table 3-1:OSA-CBM six-step architecture for PdM (OSA-CBM, 2022)

Advisory Generation (AG), Decision Making (DM)
Prognostic Assessment (PA)
Health Assessment (HA)
State detection (SD)
Data manipulation (DM)

Data acquisition (DA)

A six-step procedure for PdM applications consists of pre-processing (of data), feature extraction, anomaly detection, diagnosis, prognosis, and mitigation. The deep belief network (DBN), variational autoencoder (VAE), generative adversarial network (GAN), autoencoders (AE), and a self-organizing map are only a few examples of the deep learning approaches included in the vast literature study (SOM). State-of-the-art and excellent works in deep learning for PdM were reviewed and compared in accuracy and results, anomaly detection methods were tabulated, and conclusions were drawn for the latest work on deep learning for PdMand industrial processes (Serradilla et al., 2022, pp. 2–24).

3.2 Comparison of Machine learning algorithms for predictive maintenance

3.2.1 Supervised learning

In case of a labelled data, the data can be classified based on their labels into multiple classes. This is equivalent to applying supervised learning techniques such as regression and classification. The regression and classification are studied in this context.

❖ Regression

Linear regression is a technique for expressing the relationship between two variables using an equation.

$$h_\theta(x) = \theta_0 + \theta_1(x) \qquad (1)$$

where $h_\theta(x)$ = hypothesis

$\theta_i = value\ of\ paremeters\ to\ be\ chosen\ during\ the\ training\ process.$

x = training examples x is for input training data and y = output training data. The equation of the linear line can relate x and y, for example, slope calculations.

$$y + ax = b \qquad (2)$$

Equation (2) can be used to predict the continuous-time real-value output. taking an example of the size of a house to be predicted, a training set of data is passed through a learning algorithm (a common example of which is gradient descent) and a hypothesis function is obtained. If the size of the house is passed through this hypothesis, the model will give a predicted or estimated price of the home. Some major types of regression from an algorithm implementation point of view are

- Polynomial regression
- Logistic regression
- Lasso Regression
- Ridge regression

❖ Classification

Regression predicted continuous real-valued output; if we need to predict discrete value, classification is the preferred choice. Classification is putting data into different classes based on the similarity of features. The best example of classifications is support vector machines, which are detailed in the algorithm's examples section for their applications in predictive maintenance.

3.2.2 Unsupervised learning

❖ Clustering

Clustering is the concept of similar grouping features or similar characteristics of objects. While this is a point is under discussion, it could be argued that classification also does the same, i.e., classifying similar characteristics objects to one class, however, that is a supervised learning task and clustering is unsupervised learning, i.e., data under development is unlabeled data, and clusters are formed based on similarity. Two of the most popular techniques for clustering are K-Means and DBSCAN. K-Means was proposed posed by a bell labs employee, Stuart Lloyd, in 1957 as a method for pulse-code modulation, but it got published in 1982, without the company. The paper was titled "Least square quantization in PCM". Clustering can be applied to segmentation, semi-supervised learning, pre-processing, and anomaly detection using gaussian mixtures (Géron, 2019, pp. 264–288).

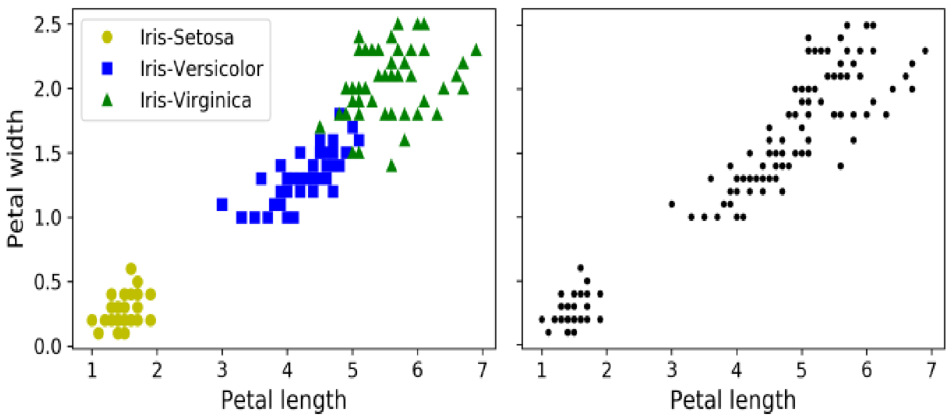

Figure 3-1: Classification (left) vs Clustering (right)(Géron, 2019)

❖ Anomaly detection

An excellent example of anomaly detection is an aircraft engine, and collecting data on parameters such as heat generated and vibration intensity. If a new aircraft engine is made operational, anomalies need to be detected, and quick decisions has to be taken if the aircraft engine needs to go for more testing. The outliers' features are flagged as an anomaly if they are less than or equal to a pre-defined threshold. Some typical applications are fraud detection, manufacturing such as predictive maintenance, and data centres for monitoring, computers, CPU load, and network traffic. The other major anomaly detection algorithms are Fast-MCD, isolation forest, Local outlier factor (LOF), and one-class SVM, as detailed in (Géron, 2019, pp. 292–299)

3.3 Algorithms examples

Deep learning and Machine learning provide a wide variety of algorithms that are suited for varied applications. In this study, only the algorithms that have industrial applications and are specifically related to PdM applications in the context of industry 4.0 are studied.

3.3.1 SVM (Support Vector Machines)

SVM is a classification algorithm that uses the concept of support vectors which are hyperplanes. For an n-dimensional space, a hyperplane is an n – 1 subspace. One-dimensional space is used to divide a two-dimensional plane, and a 2-dimensional hyperplane is used for dividing a three-dimensional subspace. This concept is similar to separating two classes by a simple linear line separating the two classes. The data is termed as linearly separable. The separating boundary is termed a decision boundary. SVM is probably the most powerful machine learning technique capable of separating linear, non-linear, and even performing outlier detection and regression. The data mostly found from sensors or any other source is not linearly separable; SVM uses the concept of kernels such as polynomial, gaussian RBF, etc. (Albon, 2018, pp. 282–293).

3.3.2 Random Forest

The random forest belongs to a class of machine learning methods referred to as ensemble methods. Supposedly, if the data contains several predictions, it will be better to aggregate the results and get a combined or aggregated result. That will improve accuracy. This is the idea behind ensemble methods. There are two types, tree-based methods, called a random forest, and boosting-based methods, termed Adaboost and XGboost. The concept of bagging is applied by aggregating the output of many classifiers (Breiman, 2001, pp. 1–10). Random forest is another powerful method that has found industry applications such as predictive maintenance. The random forest has been used for predicting relative humidity (Prihatno et al., 2021, pp. 1–3). The random forest has also been used for the Real-time PdM of wind turbines using Big Data frameworks. Big data applications and architectures have been developed by using Apache Spark (data processing) Apache Kafka (data acquisition), Apache Mesos, and Apache Zookeeper (data clusters management) (Canizo et al., 2017, pp. 1–5)

3.3.3 Recurrent Neural Networks:

Recurrent neural Networks (RNNs) are a class of networks that use previous outputs to be used as inputs and influence future outputs. This is similar to having the memory in the network to keep the information of the previous output state and use it for future outputs. There are many configurations of RNNs, such as one-to-one, one-to-many, many-to-one, and many-to-many. The primary activation functions used for RNNs are Tanh, Relu, and sigmoid. Many applications include language translation, also known as machine translation (incorporated in google translate), natural language processing (NLP, integrated as Siri, Voice search), speech recognition, sentiment classification, image captioning, as well as the creation of music. One excellent example of the application of RNNs in an industrial environment is wear prognosis and the current signature of industrial plants (Küfner et al., 2021, pp. 1–4). Deep learning techniques, including ReLu (rectified linear unit), short-time Fourier transform (STFT), and adaptive moment estimation (Adam) are used in PdM to overcome the significant obstacle of data selection. Electrical current signature is suggested as the database for this technique. Estimating remaining useful life (RUL) parameters from the preliminary data and predicting the machine failure time are two more applications of RNNs for PdMis that have been studied. (Rivas et al., 2020, pp. 1–3). There are two main types of recurrent neural networks, LSTM (Long Short-Term Memory) and Gated recurrent units (GRUs).

❖ LSTM (Long Short-Term Memory)

This techniques stem from Natural Language Processing (NLP) and deep learning and it is an artificial recurrent neural network type. This is in contrast to regular CNN (Convolutional neural network) and RNN, which do not have a feedback connection. LSTM is capable of processing image data (single point data) as well as speech data (sequence of data). An LSTM cell consists of three gated, input, output, exit, and a cell. The cell keeps information on numbers; gates are used to enter/exit information from cells. LSTM solves the vanishing gradient issue encountered while training conventional RNNs (Kane, A, et al., 2022, p. 3).

3.4 Big-Data Improvement For Predictive Maintenance

PdM using advanced data analytics enables pilots to tell the operator whether the aircraft requires maintenance or not, a big step for maintenance 4.0. The big data role is essential in this feat; big data sands on four provisions, volume, velocity, variety, and value. Primary big data services can provide four types of analysis: classification, clustering, regression, and association. With advanced big data systems and precise sensory data acquisition availability, better accuracy maintenance decisions can be made just in time (JIT). The critical aspects of PdM in Maintenance Policy Management (MPM) are criticality, availability of sensory data, reliability, timeliness, relevance, and knowledge-oriented strategy (Lee et al., 2017, pp. 4–9).

3.5 Implementation Challenges For Deep Learning in PdM

After going through the deep learning algorithms, some challenges in implementing deep learning for PdM have come across that need to be solved for algorithms implementation in the industry for PdM.

3.5.1 Lack of Sufficient Data for PdM System Creation:

The data originated from sensors; hence if sensors are new or the logging method of data puts constraints on data collection, new approaches need to be adopted for sufficient data collection for decision-making. Several processes work in a "feast or famine" mode of operation where very small or zero data is captured until a problem or fault in the machine happens. Some sensors log event codes and time stamps that tell engineers about an operation that has occurred but do not provide sensor readings at the instant of the failure. Other challenges include.

3.5.2 Lack of the Failure Data Required for Precise Predictive Results.

There are tools available in MATLAB, such as failure mode effects analysis (FMEA) that can give a good start for determining which types of failures to simulate and understanding them.

3.5.3 Understand Failures But Cannot Predict Them In Advance.

The failures can be understood but predicting it has its challenges. This problem also needs modelling in software.

3.5.4 Lack Of Sufficient Skillset For PdM Application:

These are problems in accessing data, processing the data, training of models, and successful deployment (Mathworks, 2021, pp. 1–8).

4. Predictive Analysis Techniques and case studies

The traditional control of equipment health, which acts retroactively and reactively and discovers problems after they have happened, is a typical tactic in the industry.However, an ideal PDM system must be able to predict such defects at a somewhat earlier stage, providing sufficient time for the maintenance directive afterwards and reducing the probability of failure equipment and operator downtime. Automatization of PdM (APdM) is thus emphasized for open loop control systems as the process's complexity ten folds, and manual supervision or inspection is not recommended but considered a bottleneck. A variety of case studies are presented by (Lughofer & Sayed Mouchaweh, 2019, pp. 13–23), creating three sections of an Automated PdM system.

Many methods for applying and analyzing PdM in Industry 4.0 are based on industry and manufacturing. Hence, most PdM systems are categorized as industrial types. Plants, mechanical, electrical, oil industry, and manufacturing have their analysis and application methods. The analysis methods are defined in general and later elaborated based on the industry type.

4.1 PdM Analysis Types:

There are scientific ways to gather and monitor the system information or the state of the equipment. These methods are based on the type of machinery that needs to be monitored for perfectly timed repairs. PdM can also be understood by the following examples based on each technique.

4.1.1 Infrared thermography analysis:

The temperature of the equipment is monitored, which protects the equipment from overheating.

4.1.2 Vibration analysis:

The equipment is monitored against imbalance, looseness, wear and tear, and other parameters that can lead to issues in maintenance.

4.1.3 Current analysis:

This is done to monitor the current and voltages applied to an electrical appliance, which can also be used to detect problems with electrical, magnetic couplings, and belts problems.

4.1.4 Ultrasonic Analysis:

This corresponds to the equipment's frequency change as it goes through degradation over a given period of time or usage.

4.1.5 Oil analysis:

This technique is focused on equipment lubrication level since proper lubrication is vital for the smooth operation of the asset, appliance or machine (Lisowski, 2022, p. 3). These techniques are elaborated on further in the following section.

4.2 Mechanical industry techniques

4.2.1 Vibration Analysis

In industries, most of the machines are mechanically operated. Hence a straightforward but narrow-minded approach is to conduct vibration analysis. However, additional variables and methods must be used to assess the dependability and effectiveness of machines. These include thermography, tribology, process parameters, visual inspection, ultrasonics, and other nondestructive testing methods. However, vibration analysis has excellent applications, such as gear fault diagnosis (GFD). The standard most common gear faults include Broken tooth, crack, even wear, different axis, Eccentricity, and pitch error. THESE faults can be extracted as features using deep learning methods and viewed in the frequency domain for analysis. The different techniques for adaptive mode decomposition (AMD) have been compared in the study (S. Zhang et al., 2022, pp. 1–8) for gear fault diagnosis.

In the previous ten years, the vibration parameter collection method has used single-channel data collectors based on a microprocessor and Windows-compatible software to evaluate the vibration energy of electromechanical systems. As part of the paper (Tama et al., 2022, pp. 3–6), an excellent attempt has been made to detect faults in rotating machinery using deep belief networks (DBNs), recurrent neural networks (RNNs), generative neural networks (GNNs), and graph neural networks.

4.2.2 Limitation of technology method

The more straightforward data acquisition process can result in catastrophic failures for equipment. The relatively less trained staff can result in misinterpretation of the values of sensor data.

❖ Single channel data:

The single channel of the microprocessor-based system is a limitation in the sense that if the vibration profile is to be captured constant and the speed of the motor remains within specific values (5-10 rpm for example), the results can be valid, However, other than this limit, the diagnostics accuracy is always doubtful (Mobley, 2002, pp. 99–112).

❖ Steady-state data:

Sudden changes in load or transient events are not captured with such microprocessor-based systems.

❖ Low-Frequency Response

These vibration measurement systems cannot capture low-frequency variations or low vibration speeds. Specifically, many commercially available vibrations don't function properly for vibrations measurements below 600 cycles per minute (CPM) or 10 Hz. This is due to electronic noise at low frequencies and the transducer's response dominating the vibration analysis profile, which gets lost in cumulative noise.

❖ Averaging

Every mechanical equipment is subject to random, nonrecurring vibrations and periodic oscillations. Averaging the overall signals measured over any samples can result in better results and the elimination of spurious. Averaging also improves the repeatability of the data because only the continuous signals are retained.

4.2.3 Data Analysis

4.2.4 Signal Processing analysis and techniques

4.3 Tribology

Tribology includes the completer fields of wear, friction as well as lubrication. This term is used for the design and parameter allocation of bearing-lubrication-rotor-based dynamics of machinery. In this case, light has been shed on wear and lubrication regarding oil analysis.

4.3.1 Oil and particle analysis

Oil is an essential element in the lubrication of machines in mechanical and electrical industries. Oil analysis is a technique used for accessing the content of lubricating oil, and it is not a technique directly for health monitoring of equipment or fault indication.

The primary applications for lube oil analysis are quality control, decreasing lubrication oil inventories, and estimating the perfect period for an oil change. The types of oils that can be monitored using this type of analysis continuously are Lubricating, hydraulic, and dielectric oils, followed by which the right decision for changing oils can be made.

4.3.2 Wear

Whenever a body and antibody come into interaction, wear occurs. This happens mainly due to the unavailability of lubrication. This happens due to the continuous removal of material from the surface of a solid. The types of wear that can occur are Abrasive wear, Adhesive wear, Cavitation, Corrosive wear, cutting wear, Fatigue wear, and Sliding wear. However, the four most common and essential wear types are defined here:

❖ Surface fatigue

Surface fatigue originates, from cracks, cracking growth, and separation of wear particles that are created by varying loads in the base body and counter body.

❖ Abrasion:

In this type, wear is caused by periodic ploughing and the base body fractures that are produced by the counter body asperities and hardness(Kovaříková et al., 2009, pp. 1–4).

❖ Adhesion

Assuming the adhesive mechanism's bonding force is higher than the softer friction partner's. In that instance, the material can be moved to the harder friction partner and disengaged from the deformed surface of the more relaxing friction partner. The substance has three possible states: attachment, detachment, and return to the original surface.

❖ Tribochemical reaction

This is a chemical reaction occurring during the mating of two mechanical surfaces. Therefore, the region of a chemical reaction on a mechanical surface determines its source of origin. This region can be inside the contact area (ICA), close to the contact area (NCA), such as between a sphere and a plan, or well outside the operating contact area (OCA) (Nakayama & Martin, 2006, pp. 1–4).

The best indicator of the equipment's health is wear particles. There are many techniques for determining the type of particle and their concentrations, for example.

a. Spectrometric analysis
b. Particle counting
c. Direct reading ferrography
d. Analytical ferrography

Quality oil can reduce the wear percentage of the machine to a great extent, and some parameters for oils analysis are.

a. Viscosity
b. Solid’s content
c. Water content
d. Total acid number
e. Total base number
f. Flash point

4.3.3 Limitations of Tribology

The primary and significant limitations in tribology applications for PDM are:

1. The price of equipment and machines (the capital costs and recurring costs)
2. Getting the precise oil samples
3. Extraction of results from raw data

4.4 Energy-Related Tasks and Miscellaneous Tasks

❖ Temperature Measurement

Temperature measurements are an important aspect of predictive and preventive maintenance in manufacturing. Temperature sensors can be used to monitor the temperature of equipment and machinery, providing valuable data that can be used to predict when a failure may occur.

This data can be used to schedule preventative maintenance before the failure occurs, reducing downtime and increasing the overall efficiency of the manufacturing process.

Temperature readings can be used to identify unusual temperatures that could point to a problem with the gear or equipment. For instance, a rise in temperature could be a sign that a bearing is deteriorating or that the lubrication system needs to be repaired. By keeping an eye on the temperature, it is easy to identify these problems before they become serious and schedule maintenance.

Temperature readings can also be used to identify problems with machinery and equipment cooling systems. For instance, if a machine's temperature is high all the time, it can be a sign that the cooling system is malfunctioning. It is possible to identify these problems and plan maintenance to address them before they result in a failure by keeping an eye on the temperature.Temperature readings can be used to identify unusual temperatures that could point to a problem with the gear or equipment. For instance, a rise in temperature could be a sign that a bearing is deteriorating or that the lubrication system needs to be repaired. By keeping an eye on the temperature, it is easy to identify these problems before they become serious and schedule maintenance.

Temperature readings can also be used to identify problems with machinery and equipment cooling systems. For instance, if a machine's temperature is high all the time, it can be a sign that the cooling system is malfunctioning. It is possible to identify these problems and plan maintenance to address them before they result in failure by keeping an eye on the temperature.

- Advanced Visual Techniques

In manufacturing, plants, and equipment, predictive and preventive maintenance techniques are used to foresee probable breakdowns or issues and take preventive action before they arise. These maintenance methods can be made more effective by combining them with more sophisticated visual tools, such as computer vision. Computer vision can be used, for instance, to track and assess the state of machinery and equipment, spot possible issues, and notify maintenance staff to take appropriate action before a breakdown takes place. In order to find possible problems before they result in failure, visual inspection techniques can also be used to do non-destructive testing on components, such as welds. Overall, the effectiveness and dependability of industrial operations can be increased by the application of advanced visual techniques in conjunction with predictive and preventive maintenance.

4.5 Other PdM Techniques

Some other PdM techniques include Ultrasound and infrared thermography, and the following few lines touch on this topic.

4.5.1 Ultrasound

Ultrasonics technology is versatile, easy to use, and cheap to implement. This technology is considered a major PdM technology when combined with vibration analysis and infrared thermography.

❖ Ultrasonic applications

The primary applications for ultrasound technology are compressed air leak management, condition-based monitoring of equipment (UCM), condition-based lubrication for acoustics, electrical applications (switchgear, grid stations, and high voltage transmission and distribution lines), reciprocating compressors and valves, cavitation pumps, heat exchanger and condenser leaks, valves and hydraulic leaks (Rienstra, n.d., pp. 1-3).

❖ Limitations

In manufacturing, ultrasound and ultrasonics are frequently utilised for a range of tasks like welding, cutting, cleaning, and measuring. Ultrasound and ultrasonics have some restrictions when used in manufacturing, though.

One drawback is that some materials, such as brittle materials that are prone to breaking or metals with limited thermal conductivity, are not suitable for use with ultrasound or ultrasonic technology. Additionally, certain materials can absorb or scatter ultrasound and ultrasonic waves, which can reduce their effectiveness.

Another drawback is that the precision of measurements and the efficiency of cleaning and welding operations can both be impacted by ambient conditions that affect ultrasound and ultrasonics, such as temperature and humidity.

Additionally, the heat produced by ultrasonics can cause the material to expand and shrink thermally, which could have a negative impact on the finished product.

Last but not least, certain firms may find it difficult and expensive to deploy and maintain ultrasound and ultrasonics.

In general, when choosing whether to utilise ultrasound and ultrasonics, it's crucial to take into account the particular requirements of a manufacturing process and the qualities of the materials being employed.

4.5.2 Infrared thermography

Thermography idea generates from the working operation of infrared cameras. When this idea is applied to machines, it can predict the areas with more thermal energy than the surroundings or the temperature than it should be, the so-called concept of thermal anomalies. IR cameras are designed to detect humans based on their emitted energy and the temperature difference between the body and surrounding ambient temperature. The trained technician can learn a lot from this information about the machine's current working dynamics and status. Infrared thermometers, line scanners, and line infrared imaging cameras are examples of further infrared systems.The major applications of IR thermography are in electrical equipment, mechanical equipment, and energy systems (Mobley, 2002, pp. 118–121).

4.6 Failure Mode Analysis

Failure mode analysis is a set of techniques used to assess the importance of the components that need PdM the most to improve the system's overall efficiency.

Reliability analysts have used many techniques to measure up the importance of the equipment for the pant and adjust the maintenance planning, such as Root cause analysis (RCA), failure mode effect analysis (FMEA), and fuzzy methodology (FM). Some other methods are reliability block diagrams (RBDs), Monte Carlo simulation (MCS), Markov modelling (MM), fault tree analysis (FTA), and Petri nets (PN). In these methods, the issues relating to the system's unreliability can be estimated using RCA and FMEA. This uncertainty present and detected hindering the system's performance can be modelled using fuzzy logic and modelling techniques (Sharma & Sharma, 2010).

4.6.1 Failure Mode Effect And Criticality Analysis (FMECA).

A conventional PdM setup will monitor every piece of equipment for performance monitoring; however, a complicated production system cannot monitor every piece of equipment. FMECA is a semi-quantitative technique that identifies key equipment and aids in prioritising the equipment so that PdM can be applied to enhance system performance. The FMECA serves as a starting point for reliability engineers to provide an estimation matrix made up of parameters like the failure rate or force of mortality (FOM), mean time to repair (MTTR), mean time to failure (MTTF), mean time between repair (MTBR), and preferably some other parameters as well. A plan for PdM development follows the diagnosis of the necessary PdM equipment.(Chukwuekwe, 2016, pp. 37–38).

The FMECA is a system analysis tool that can function as input for logistic support analysis, test requirements document (TRD), fault trees, built-in test (BIT) analysis, and safety analysis. FMECA is now used in various studies, such as BIT analysis, LSA-B sheets, fault trees, and TRD that may be automated for total quality management goals. Numerous strategies were created and presented to improve the effectiveness of FMECA analysis and its documentation. (Luthra, 1991, pp. 1–4). A good maintenance strategy is specific to a particular industry. A modified FMECA is developed under the following steps for strategy implementation.

1. Tabulating the system under consideration, subsystems and other components of the system need to be included in the analysis;

2. Tabulation and description of failure modes for every part and sub-part of the system;

3. Predicting the probability of failure (Sf) for each failure mode listed in step2, the likelihood of missing detection of failure (Sd), and the rank of criticality and severity of the losses (S);

4. assigning a criticality number as per risk level for each fault (risk priority number, RPN= Sf*Sd*S);

5. Reordering the faults and issuing according to the RPN;

6. prompt action towards highly critical and risk-producing faults;

7. A revised risk analysis verifying the effectiveness of the actions taken. The Following factors need to be taken into account.

Safety: Equipment importance for each process; The budget and price of maintenance; the frequency of faults, Downtime period; Operating conditions; Machine access difficulty (Bevilacqua et al., 2000, pp. 3–7).

The tree analysis is also used for FMECA Modelling. The conventional fault tree analysis has been proven to be o be ineffective in dealing with the imperfection of input failure data and the uncertainties involved in the modelling process, considering especially the dependency of failures. A new Fuzzy Fault Tree Analysis (FFTA) is proposed based on the probability distributions associated with primary events and summing these events using fuzzy logic and algebra theory.

A proposed modelled equivalent representation for conditional possibility distributions is employed in combination with AND and OR-gates of the fault tree to evaluate the dependency of failures (Misra & Weber, 1989, pp. 1–4).

The practice of ranking and reducing the effects of non-valued added effects in manufacturing is vital to applying sustainability in manufacturing. The research-oriented approach towards ranking the risk of non-value-added activities is focused on the product design and manufacturing sectors. This practice is rare in maintenance engineering. A modified FMEA (Failure Mode and Effect Analysis) can be devised to estimate the criticality of the maintenance sector's waste. To help experts assess the criticality of garbage, a modified and improved model for ranking the risk of maintenance waste using the Waste Priority Number (WPN) is proposed (Sutrisno et al., 2015, pp. 1–5). Another attempt deals with the case of hydraulic turbines as a case study and implements the FMEA and FTA analysis for employing reliability-centred maintenance. Concerning the reliability-centred maintenance (RCM) specifications, FMEA (Failure Modes and Effects Analysis) and FTA (Fault Tree Analysis) in hydraulic turbines are modelled. The FTA and FMEA methods for product design, equipment, and processes analysis to perform the systematic and standardized estimation of possible failures for controlling the implementation of corrective or preventive actions and documenting the consequences. The priority risk number (PRN) is a metric produced by FMEA for ranking and prioritising critical failure modes in the system. A threshold such as PRN higher than thirty declared that component to be in a primary failure mode that requires immediate action. The FMEA analysis applied in a plant test case resulted in two crucial elements in the lubrication and cooling system of the combined bearing: the filter 01 and the motor pumps' command and control circuits. The results showed that heat exchangers can present more than twenty (20) faults and are still not considered critical components since these faults have insignificant effects on the operation and are easily reparable(de Queiroz Souza & Alvares, 2007, pp. 8–10).

A detailed study in the form of a state-of-the-art review is presented after studying a large group of 202 scientific papers searched using keywords and 109 patents. Excellent classification is done based on literature sources (consisting of a major journal, proceedings, and conferences) and classification based on problem classes (such as applicability, cause and effect, risk analysis, and problem-solving). The paper presented the summary of the problem statement and ended up with suggestions for ameliorating and improving the FEMA (Spreafico et al., 2017, pp. 3–7).

Finally, a new standard is proposed that seamlessly integrates PdM into modern product design practices by implementing three significant changes.

(1) This standard mandates FMECA procedure as complementary and to be implemented from the start to finish of the product development cycle, not just as a process to be implemented after design completion, by focusing on functional and interface FMECAs as that of the traditional piece part FMECA.

(2) The concept of "failure mode equivalence" assists in minimising the extra work required for equipment-by-equipment fault and failure analyses; this is also integrated into the design process from the outset to be carried over to the effects of interface and piece-part failure modes later in the design.

(3) A Pareto ranking method based on probability and the severity of the associated failure mode is used to rank criticality. Unlike the earlier convention of criticality numbers described in Mil-Std-1629, this measure is more easily applied to various practical applications. The mathematical difficulties of the RPN analysis utilised in the Automobile FMECA standard, SAE J1739, are overcome by this criticality assignment approach (Bowles, 1998, pp. 1–3).

4.7 Establishing A PdM System

Many PdM program efforts during the last three years have unfortunately been unsuccessful since there was no clear objective and motivation. Apart from the initial setup, an enormous amount of operational costs is required. A PdM programme needs to set its long-term objectives, some of which include:

1. It is vital to stop unnecessary maintenance.
2. Less loss of output due to equipment malfunction.
3. Reducing the repair part inventory is necessary, which will increase process effectiveness.
4. The plant's lifetime of operation and product standards should be raised to enhance capacity.
5. Revenue increases and a decline in net maintenance expenses. The maintenance must be stopped, which is not essential.

The functional requirements include management support, detail-oriented, accountable personnel, efficient gathering of data and analysis procedures, and a highly reconfigurable database stage.

A vibration monitoring system is most challenging to establish and has a tremendous return on investment. It is also the most used technique in most PdM programs. While selecting a PdM system, the following considerations must be met and considered.

- Easy-to-use software and hardware
- Automated data acquisition
- data management,
- Trending flexible, highly reliable, and accurate system
- Easy access to training and technical support.

4.8 Benefits of the PdMsystem

PdM can be used as a tool for any of the following, based on the use cases and techniques used:

- As a maintenance management tool
- As a plant optimization tool
- As a reliability improvement tool

Multiple surveys have attempted to derive the benefits of PdM. Almost 91 per cent (90.9%) of companies participating in surveys reported measurable savings by adopting an intelligent PdM program. According to reports, maintenance costs and downtime have recovered by an estimated 113%, an improvement of 13% over the prior system. With an average budget of ($12,053,000) for each company as a participant, each participant saved $1.6 million. A triumphant return on investment should be 10:1 and 12:1 for a perfect implementation of the PdM program. Another statistic is that the production should benefit from $10 to $12 for every dollar invested. This was not the case looking at the survey results. In practice, the average (ROI) return on investment was only 1.13:1. This will make many managers dubious about implementing PdM.

The statistics reported by the survey might be misdirecting. Carefully observing, it is evident that 26.2 per cent of participants answered their PdM implementation recovered the invested costs; 13 per cent did not have numbers, and 50.8 per cent failed to get back investments. This brings into question if the PdM is worth the effort, and it is answered by 10 per cent of respondents of the survey. Fifty per cent of the factories reported profits five times greater than the entire set-up expenses invested, or a 5:1 return on investment, demonstrating that this 10% of enterprises not only paid back the investments but also generated a reasonable profit. It has a significant influence and steers the business in the proper direction, even though this still compares favourably to other successful PdM deployments. The statistics also confirm that few facilities fully utilise PdM's potential. When fully utilised, these technologies could generate returns on investment that are significantly higher than 100:1, or $100 for every dollar invested. As we have often emphasised, technology is readily available, but in order to fully benefit from it, it must be used appropriately. The poll results unmistakably demonstrate that many businesses are still experiencing this (Mobley, 2002, pp. 61–70).

4.9 Case Studies

This section describes three case studies of PdM applied in practice; the first actual life application is PdM in elevators maintenance, plant maintenance and PdM in manufacturing.

4.9.1 Cast study 1: Predictive Elevators Maintenance

As the industry continues to develop, many elevator productions and repair service enterprise has sought to upgrade their service policy from preventive to predictive maintenance. LoTand cyber-physical systems, data mining, and intelligent elevators are the trends making the concept of PdMservice policy by remote monitoring and fault detection when needed. A systematic study of the Chinese elevator industry, growth rate, maintenance, and repair has been presented, and an intelligent PdM system is proposed for automatic remote monitoring and fault detection as well as the safety aspects of the elevator maintenance industry. An Intelligent Faults Diagnosis and Prognosis System (IFDaPS) is proposed in industry 4.0, challenging the conventional term PdMas PdM cannot be implemented without intelligence and hence a more practical, automatic approach named intelligent predictive maintenance. The integral parts of industry 4.0 are described as Cyber-Physical Systems (CPS), IoT, Internet-of-Service (IoS), and Data Mining (DM)(K. Wang et al., 2016, pp. 3–6).

Commented [AU1]: How a word can be challanged? This paper is not about text analysis.

Commented [HH2R1]:

Commented [HH3R1]:

The maintenance department for elevators can quickly get negligent for routine maintenance; nobody cares unless a lift gets non-operational. This situation can result in human and capital loss, or injury, if not corrected in the long run. Maintenance and safety are the two most important issues. It is necessary to create a strategy and parameters that may not only detect impending failure but also give a general picture of the complicated problems encountered in both predictive and preventative elevator maintenance. Although feature engineering has been explored, its applicability has practical limitations, including its situation, data specificity, and difficulty generalising to all applications. Deep learning's random forest, which has an accuracy of 91.50

percent when applied to an open-source dataset, is more valuable (Awatramani et al., 2022, pp. 1–3).

Elevators and machine tools are two entirely different applications that fall in the category of non-critical machinery. Elevator monitoring and maintenance is a novel technology and does not have a specific skill set or previous experience for monitoring. As illustrated in section 4.2.1, the machines case study is based on vibration analysis, and there are data and experienced personnel available in the vibration analysis. The neural networks are trained in cases where there is no previous data or professional staff available for making predictions on repair and detection times. Bayesian networks have been implemented as another approach with the ability to model knowledge in the network. A sensor processing unit and remote monitoring system complete an automated remote condition monitoring system for test applications such as elevators and machine tools (Gilabert & Arnaiz, 2006, pp. 1–4).

Among all the techniques for PdManalysis, oil lubricating analysis can be a perfect choice for hydraulic lift applications. There is an observation that lifts maintenance staff does not always follow the manufacturer's recommendations. Hence, the methods are recommended for oil analysis, percentage of the metals, corrosion, and the wear resistance estimation of the connecting mobile parts. Power transmission (mainly used for lifting work) in hydraulic systems happens through pressurized liquids. Two primary hydraulic fluids are mineral oils and synthetics, and this classification is based on viscosity grade. Working lubricating oils in machinery is made up of complex mixtures of hydrocarbons having molar masses in the range of 250–1,000 (Kalligeros, 2013, pp. 1–4).

The IoT-based system comes to the rescue in the event of a lift issue. A good example is Singapore City, where occupants of tall buildings and sky scrapper dwellers are impacted by faults and inconveniences like worn sheaves in lift systems more than 80% of the time. The leading cause of these problems in the elevators is the old elevator system of buildings, preventive (maintenance after a fixed period relying on statistics) or reactive (maintenance after faults has happened). IoT Sensors have been installed in high-rise buildings, and data is collected from sensors, as well as the complaints of residents, suggestions of engineers, fault records of 21 types, and static data (age, the brand of an elevator, and the number of residents in buildings. A problem formulation and data pipeline are formed, and results of different random forest variants are collected with their implemented accuracies (Ma et al., 2021, pp. 4–18). The results are tabulated in Table 4.1.

Commented [AU4]: Where does this chapter provide this?

Commented [HH5R4]: resolved.

Commented [HH6R4]:

Commented [AU7]: Is this relevant for elevators? Please write this information into the sentences and back it up by a academic reference

Commented [HH8R7]: It means if we don't have experience available, we can use data and train neural networks. alternative approach.

Commented [AU9]: Why only for a test application? What about real life?

Commented [HH10R9]: These applications are what the author researched on and deep learning for predictive maintenance topic does not include real life applications? what type of real-life applications. elevators and machine analysis are real life.

Commented [HH11R9]:

Commented [AU12]: Please use a reference. Can there also be a non-perfect choice?

If you use this word like „a lot", „Strong", „perfect" etc. you need to quantify them. Like what means perfect? 99980 times of 100000 times, or 850 times out of 1000 times, and of course this need to be proofen by a academic reverence.

In case the source use such a phrase, then it is not a peer-reviewed article. Because in a peer-reviewed article this flaws do not happen.

Commented [AU13]: by whom are they recommended?

Commented [HH14R13]: The reference is at the end of paragraph, but all major books on oil analysis includes sam

Commented [HH15R13]:

Commented [AU16]: Please use a source.

Commented [HH17R16]: It is one source, for the three sentences.at the end of paragraph

Commented [HH18R16]:

Commented [AU19]: Great!!!!

Commented [AU20]: Main cause of what?

Commented [AU21]: Here you use again the abbr. IoT

Commented [AU22]: I did not see this in our paper. Why do you write „in this study" if you mean your source?

Commented [HH23R22]: 21 types of studied are many times in papers for instance on page 18..before that as we

Commented [HH24R22]:

Table 4-1: Random Forest variations and their respective performance comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Random forest	97.3	NaN	0.0	NaN
Random forest + random under sampling	59.2	4.9	77.8	9.21
Random forest + random oversampling	96.2	0	0.0	NaN
Random forest + SMOTE	94.4	14.3	22.2	17.4
Random forest + SMOTEEN	87.3	7.5	33.3	12.2
Balanced Random Forest	61.4	4.48	66.7`	8.39

Another exciting application of IoT in PdM for elevators comes from the necessary shift from CM and TBM to CBM using IoT sensors. Theoretical results indicate that CBM is better than CM and TBM but requires periodic monitoring using state-of-the-art IoT technologies. Theoretically, it is all perfect for CBM; however, the challenges lie in data collection, analytics, and decision-making. The author believed the advancement of IoT technologies could help implement remote diagnosis, fault finding, troubleshooting, and maintenance of elevators, enhancing safety and economic efficiency. The IoT sensors employed in this paper referenced at the end of the paragraph can gather in real-time 200 essential properties of elevators and more than 100 properties for the escalators. The collected parameters are inclusive of door operations, numbers of starts, stopping accuracy, operating direction (up or down), vibration and noise, braking distance, step chain speed and tension (escalator), deceleration time, and many mechanical and engineering statistics; the critical parameters measured shown in Figure 6 for elevators and escalators, respectively(Lai et al., 2019, p. 13).

Commented [AU25]: ??? You explained this already. Only use IoT

Commented [AU26]: ??? You explained this already. Only use CM

Commented [AU27]: ??? You explained this already. Only use TBM

Commented [AU28]: No!!! This is not in our paper. Again this is a pure copy and paste from you source....

Commented [HH29R28]: This is in the source; you will have to download the full paper and this information is on page 13-14. If you use Turnitin, you will find the paragraph is rephrased as much as possible, since the words have no replacement such as 200 etc.

Commented [HH30R28]:

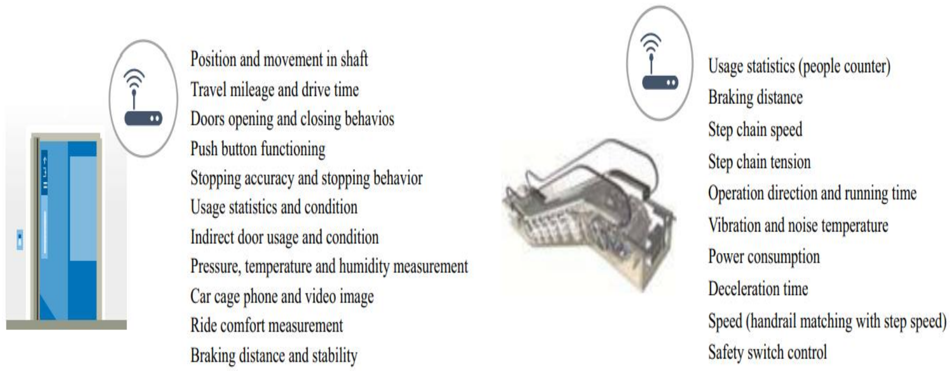

Figure 4-1: Collected parameters for elevators and escalators, respectively.

Each elevator has a capable data transmission unit (DTU) linked to the control centre, gathers data in real-time, and sends it to cloud service up to 200 times per second in three directions: A WLAN (Wireless Local Area Network) or LAN is provided to the installation: the data transmission between the cloud and installation site can take via 2G/3G/4G/5G cellular network offering Access point, in case, neither LAN/ cellular network is available, a sub-district network is introduced as a last resort.

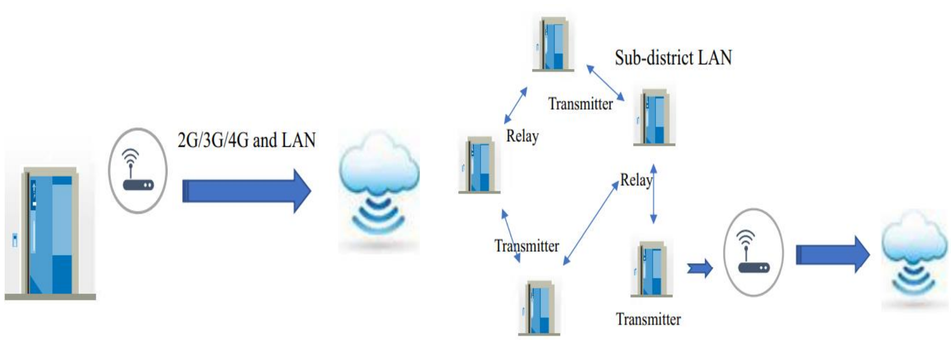

Figure 4-2: a methodology for data collection and transmission to the cloud

Kone is a world leader in elevator production and escalators installation. To face the challenges of real-time predictive analytics of elevators and decision-making, Kone collaborated with VTT, a Finnish research centre, and Prima Power, which provides services in data collection, analysis, and processing solutions in PdM for elevators. VTT did the Research and Development (R & D) part, installed the sensors at the Kone premise, and implemented the algorithms for the calculation of remaining useful life (RUL) for conveyor bearings. Prima Power aided the research with a remote cloud environment for the collection of vibration-based measurement data and Serena Customer Web analytics for condition-based maintenance(Boldosova et al., 2021, pp. 1–3).

A testbed is demonstrated starting from sensor selection and data collection, microcontroller choice, and the processing in MATLAB and deep learning algorithm implementation. A DC motor is selected to resemble the motor working in elevators highly. Two loT devices /sensors, LM35 (for temperature data collection) and Encoder Sensor Module (RPM capturing of the DC motor), were chosen for the collection of the motor parameters. The data is passed through Arduino Uno for its conversion to digital form from raw sensor data. As a system from Parallax, Inc, PLX-DAQ is used to convert this data into excel form; in this case, a comma-separated (CSV) file can also be created. Finally, MATLAB-enhanced capabilities in deep learning, interfaces, and libraries are used for Fine KNN implementation to predict desired parameters in KNN is an excellent choice after a thorough literature review (Shen et al., 2021, pp. 2–7).

4.9.2 Case study 2: Applying Predictive and Preventive Maintenance (PPM) in Plant Maintenance

The following point and areas need to be understood for good PPM practices.

- Failure Modes, Effects, and Critical Analysis (FMECA)
- Root Cause Failure Analysis (RCFA)
- The creation of machines maintenance plans
- Breakdown maintenance (BM)
- Preventive Maintenance (PM) and PdManalysis
- Reliability Centered Maintenance (RCM)
- Equipment design modification (Tshabuse & Pretorius, 2013, p. 14)

Sensors are an integral part of CBM. In plants, condition-based maintenance can be done with the help of many types of sensors; one classification of industrial sensor types can be of three types which are:

Process sensors: Existing, i.e., RTD (resistance temperature detectors), thermocouple, and pressure transmitters to get the current state of performance of sensors, also the senor-to-process junction/interface, and even the process parameters.

2nd class includes test sensors installed on plane rotatory machines and helps get vibration measurements and translate them to RUL calculation. It also has wireless sensors for other parameter collections, such as temperature, pressure, and humidity, apart from vibration amplitude.

The third method is the insertion of a test signal to measure the response of the equipment, sensor, or cables. Cable faults or responses can be measured by a technique called time domain reflectometry (TDR) (Hashemian, 2011, pp. 1–4).

4.9.3 Case study 3: PdM in manufacturing.

A case study that employs Explainable AI (Artificial Intelligence) in practice can best describe PdMin manufacturing. DARPA created Explainable AI (XAI) with two main objectives, i.e. to create more human models and explainable ML models while producing high prediction accuracy. By explainable, it is meant the models that are more readily approachable and understandable by humans and trustable and manageable by humans as created by official artificial intelligence model creators. The models such as logistic regression (LR) can be explained to management and executives; however, when dealing with deep learning and Machine learning models that are trained on data, explainability is an impossible task and an issue. Justification for decision-making by AI is a crucial need for current services and solutions. In this study (Hrnjica & Softic, 2020, pp. 3–6), The data set is collected using the Azure setup (blob storage and Azure Gallery) and classified into five classes.

- Telemetry –data about the history of equipment behaviour (voltage, current, vibration, etc.).
- Errors – The warning and errors in the equipment are error data.
- Maint – The maintenance and possible repair/replacement of machine data.
- Machines – Parameters describing the machine characteristics.
- Failures – Parameter collected when a particular piece of equipment is a result of device failure.

Different companies are studied on a case basis, and the PdM applications of these six different and their collaborator companies are surveyed.

❖ BASF / Schneider Electric

BASF, a leading chemical solutions provider, and Schneider Electric, a leading electrical power distribution supplier, are implementing digitization as a core strategy. The BASF's Beaumont, TX, plant needed to expand, and it consulted Schneider electric for deploying lloT solution for the health monitoring of equipment.

❖ ALCOA / Senseye

Likewise, Senseye implemented the cloud and sensor collection data setup for ALCOA, using Oracle softPI and Oracle eAM solution.

❖ Duke Energy Renewables / Seeq

Advanced analytics and machine learning were used by Duke Energy Renewables, a wind and solar energy service provider in the United States, to profile wind turbines and identify subpar contractors automatically. Over 2.65 years of data, a Seeq automated profiling tool was trained, producing promising results for detecting 12 errors. The company intended to apply these findings to the remaining turbine locations.

❖ Global Mining Company / Uptake

A mining company got the solution of Uptake, and the provided AI software solution is monitoring four telematics data sources for predictions of faulty wheels, wheel bearings, and axles on 10,000 rail cars.

❖ Cement Plant / Aspentech

Another cement plant chose Aspentech solution for forewarning of cyclone blockages. The previous blockage was used to train the agent and it was provided failure signatures that were left after degradation, breakdowns, and process disruptions, and the model got trained to recognize the pattern.

❖ Wärtsilä / Pega

A Finnish provider of energy and market services worked with Pega to increase its loT capability for the deployment of digitization. Wärtsilä can aggregate all of the sensor data it has collected into a cloud data lake using Pega's IoT-powered services. The data is then pre-processed (normalised) and sent to the Pega PdM setup. The data is analysed to find exceptions that might point to a growing risk of engine failure or a decline in equipment performance using a complicated rules system. (Kennedy, 2020, pp. 1–5).

In terms of PdM, a survey has been conducted on the technologies available including, planning and scheduling, anomaly detection, explanatory data analysis (EDA), and explainable artificial

Intelligence (XAI) along with a systematic literature review. Out of these four technologies, Anomaly detection is the most effective and mainly deployed solution for PdM.

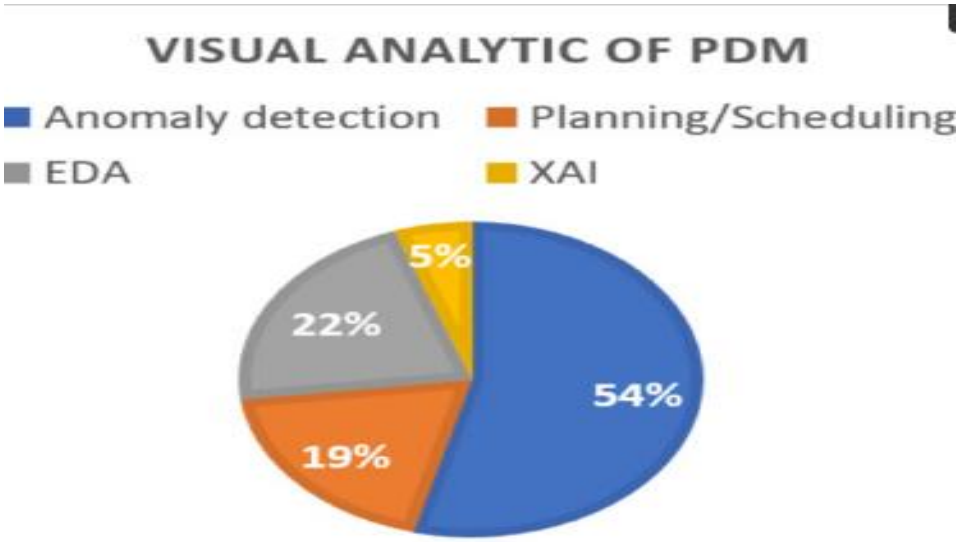

Figure 4-3:Deployment percentage comparison of four PdMtechniques (Cheng et al. 2022)

Most of the PdM studies required high-quality feature engineering. However, it is not easy to produce high-quality feature engineering, and the best method is semi-supervised learning for automation of the feature engineering (Cheng et al. 2022, pp., 3-6).

Edge computing is another technology that can be used for PdMdata transfer from sensors to the processor, in this case, edge processor. Rigorous literature research and text-mining analysis were carried out in this work. The main advantage of using Edge computing over conventional methods is the low latency for large-sized data transfer generated from thousands of sensors. The edge processor can provide immediate feedback that can be incorporated for fast decision-making in no time. However, it requires investment in infrastructure establishment as well as employee training and knowledge upskilling to make them able to handle the complexities of edge computing (Kubik et al., 2022, pp., 1-3). The leading technologies compared are Edge computing (EC), IoT, IIoT, Cyber-physical systems (CPS), cloud computing, edge and fog computing, computational offloading, blockchain, smart manufacturing, smart manufacturing, AI/ML, deep learning, game theory, digital twin and SDN.

As sustainability interest grows, the prediction of energy consumption in plants is another emerging and exciting topic for researchers. Different models for energy consumption need to be developed. One such modelling attempt is to classify machining processes into three classes: cutting energy, machining system energy, and machining process energy. The class energy category includes net cutting-specific energy, spindle-specific energy, and machine tool energy consumption in the cutting process.

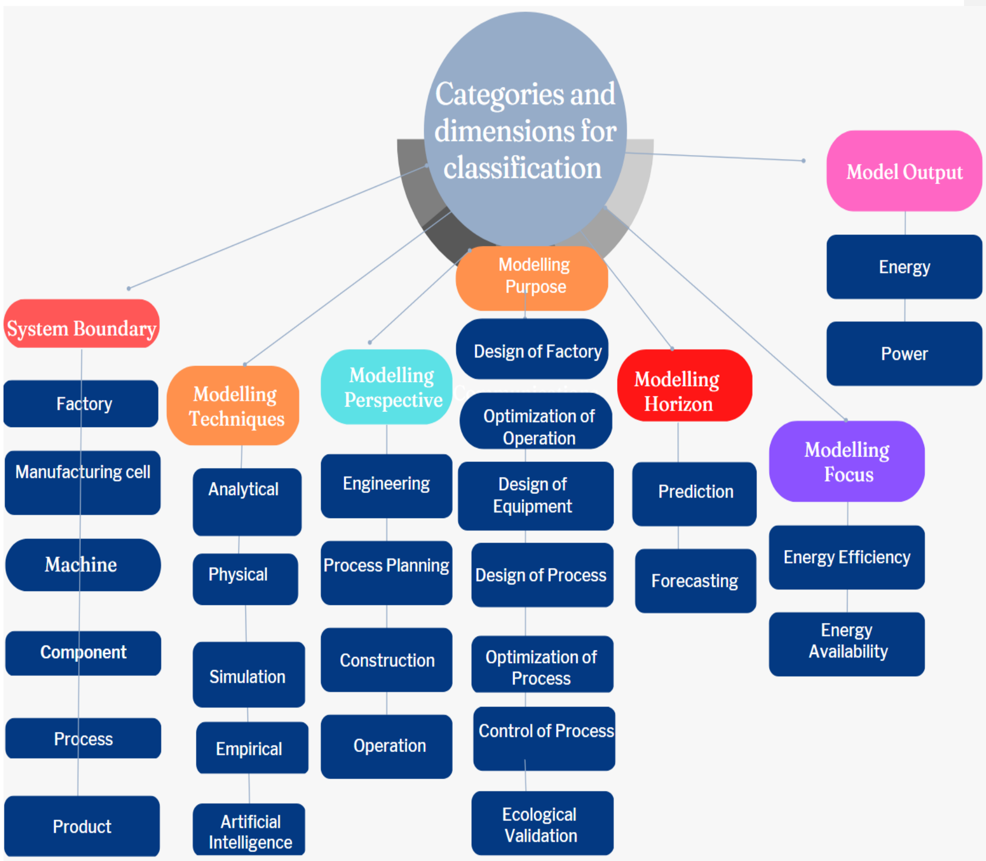

Figure 4-4: Modelling energy consumption and prediction for manufacturing plants (Jürgensen & Weigold, 2021)

This energy modelling has been done based on the system boundary, modelling techniques, modelling focus, modelling horizon, modelling perspective modelling purpose, and model output. This constitutes a detailed modelling, analysis, and synthesis breakdown followed by an overview of Machine learning, deep learning, and artificial intelligence techniques to be implemented for these energy models that are created and will be implemented for PdMat energy and manufacturing plants(Jürgensen & Weigold, 2021, pp. 2–6).

The cloud provides a service-based architecture and deploys and manages PdMservices, including machinery condition monitoring, data analysis and diagnosis, prognosis, and maintenance planning and scheduling through the internet. It is of great interest for businesses to implement cloud-based PdMservices to provide reliability, availability, and safety of cloud manufacturing. Before cloud-based PdM can be successfully deployed, a number of issues must be solved, including the development of automated service provisioning, energy-efficient data centres, traffic management and analysis, storage technologies and data management, and unique cloud designs.A novel cloud-based PdM setup based on mobile agent technology is suggested to overcome these issues and can intelligently allocate resources using embedded cloud sensing and computation nodes (J. Wang et al., 2017, pp. 1–4).

In deep learning, ensemble learning has been defined as an essential strategy to implement PdMsetups in manufacturing. PdM is useful in device management, facility management, total quality management, and equipment maintenance-related fields. Applying lloT technology to manufacturing gives birth to PdM for the manufacturing process. Cheveron corporation is a famous energy company based in San Ramon, CA, USA that deployed lloT for the prediction of corrosion and pipeline damages. The effectiveness of the big-data collection process from sensors, as well as the processing and result extraction processes using edge/cloud analytics, are improved using a novel strategy called "ensemble learning," which makes use of an adaptive boosted decision tree, which combines a neural network and a boosted decision tree, (Hung, 2021, pp. 2–8).

Two case studies were performed. They are tabulated, along with the results, PdM accuracy and performance.

Table 4-2:Case studies for PdM in manufacturing using ensemble methods

Predictive Problem	**Semiconductor Case**	**Blister Packing Machine Case**
	Yield failure in the semiconductor manufacturing process.	The quality of scissor product packaging in the blister packing machining process.

Results of semiconductor case				
Type		Ensemble	Single	
Algorithm	**Proposed Method**	**Boosted Decision**	**Decision Jungle**	**Jungle tree**
Accuracy	0.974	0.966	0.941	0.952
Recall Rate	0.957	0.945	0.925	0.928
Results of blister packing machine case				
Type		Ensemble	Single	
Algorithm	**Proposed Method**	**Boosted Decision**	**Decision Jungle**	**Jungle tree**
Accuracy	0.992	0.991	0.992	0.987
Recall Rate	0.997	0.993	0.992	1

Industry 4.0 is the core of modern manufacturing system design. A new model called PMMI 4.0, a PdM model for Industry 4.0, has been proposed in light of the significance of modelling the creation of cutting-edge technologies. It uses LSTM to calculate the machine's remaining useful life (RUL) and a novel suggested model called PMS4MMC to install an optimised maintenance schedule plan for various pieces of equipment on the plant. After presenting the theoretical foundation for PdM, and LSTM, each modelling step has been performed, including prediction and big data in the context of industry 4.0 and five different algorithms (pseudo codes) are presented, and two detailed scenarios in the context of manufacturing setup, are presented with the results and improvements obtained(Sang et al., 2021, pp. 3–14).

The main methods that have been used for PdMsetup are statistical, machine learning, and deep learning model. The statistical models necessitate significant knowledge of the domain and don't converge and fit well for highly complex manufacturing environments, and deep learning and machine learning necessitate vast amounts of data for learning and training models and their accuracy and errors are a concern after the model training, including trust issues. Besides, there is noise in industrial equipment that hampers the accuracy of measurements. Hierarchical Temporal Memory (HTM), from the neuro Science branch, is suggested by employing binary sparse distributed representations (SDRs) to represent data as well as a model incorporating feed-forward, lateral, and feedback connections; HTMs mimic the interlinkages between pyramidal neurons in the neocortex. HTMs have the advantages of being an online learning algorithm, requiring less tuning for a particular application, handling noisy data effectively, and being adaptable to changes in the data as they are periodically taught from it. This shows that HTMs can learn or train from a single training pass with minimal to no hyperparameter adjustment and still offer higher efficiency with small training datasets.(Malawade et al., 2021, pp. 1–3).

Another five-step architecture for PdM using LoT (IoT) devices is proposed starting from sensor data collection in the form of a digital signal, followed by data storage and transfer and edge/fog/cloud processing, and finally, failure prediction. The PdM system is to be implemented in a heat exchanger unit, where the primary functionality is to provide cooling for extremely high-temperature synthetic fluids getting discharged from the assembly line. The plant setup had repeated problems and downtime due to the clogging of its conduits as a result of chemical deposits; another problem was heat being leaked due to thermal cracks or perforations. This is a critical plant setup with the safety of employees at stake. The model performance on the dataset is given below in table 4-3 (Nangia et al., 2020, pp. 2–5).

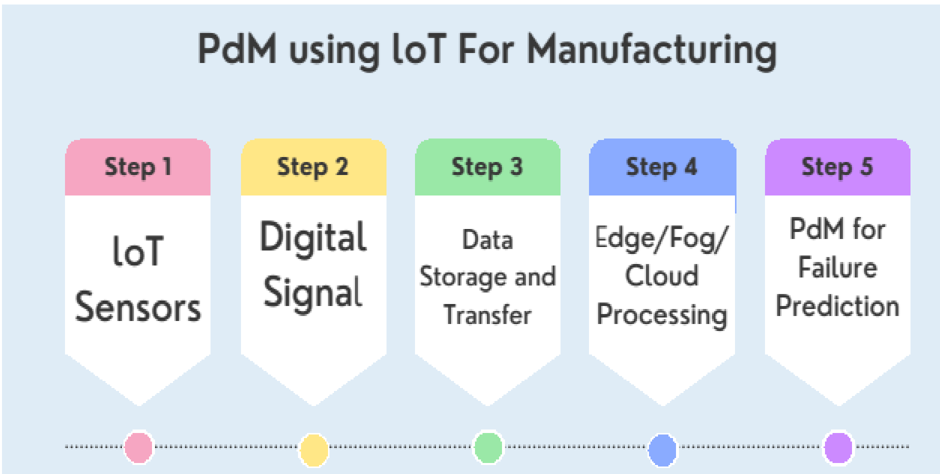

Figure 4-5:Architectutre for IoT for PdM in Manufacturing

Table 4-3: Machine learning methods and IoT for PdMin manufacturing

Metric	**C&RT**	**Boosted Classification Trees**	**SVM**
Precision	0.891	0.899	0.893
Recall	0.914	0.908	0.894
F1 Score	0.903	0.903	0.893
Error Rate	0.099	0.097	0.106

Data mining is another excellent technique for fault diagnosis and failure prediction in manufacturing. Another algorithm studied is derived from data mining and semantics, and the chronicle mining concept from data mining has been exploited to forecast the failure of the industrial equipment in advance based on monitoring and condition assessment. A Manufacturing PdMOntology (MPMO) has been exploited to forecast or predict temporal constraints of failures using its rule-based extension, and prediction results are presented formally. From Semantic Web Rule Language (SWRL), a set of rules is created to predict the time of equipment failures in advance accurately (Thakker et al., 2020, pp. 1–5).

Catraport, Lda, a company established in Portugal in 2015, is aimed at producing components, spare parts, and accessories for the automobile industry using the procedure of cold industrial stamping. The main activities are cold stamping, plastic stamping, creating tools and moulds, welding, assembly, and painting components. For system synchronicity and to keep the production line automatic, many Sinamics motor controllers and a Siemens PLC S7 series are used. For this production setup, an online monitoring tool consisting of three tools 1) a visualization tool 2) a monitoring tool; the first and second tools are developed on platform Node-RED (for the implementation and development of rules), and the third is 3) a mobile Android application created on the MIT App Inventor platform. The whole online monitoring solution is according to Nelson's rules. The system can sound alarms when a potential disturbance or abnormality is anticipated and occurs after monitoring the equipment's state. Additionally, an Android application was created that enables any staff member to rapidly confirm the system state outside of the shop floor, allowing the maintenance professional to examine the machine's status (Cachada', 2018, pp. 20–84).

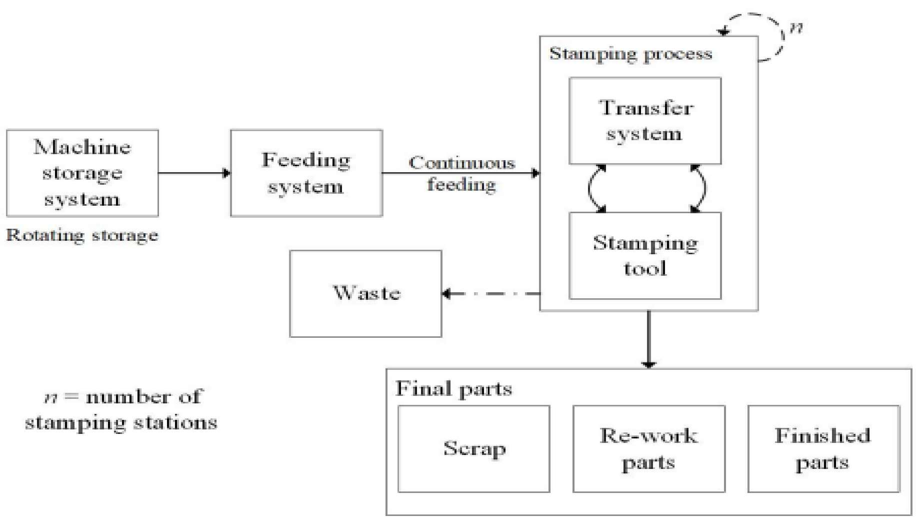

Figure 4-6:Stamp collection procedure in Catraport, Lda

The last paper study used digital twins and physics-based simulation models, using prognostics and health management techniques in manufacturing scenarios. A formal literature review is conducted in the first stage, with the same objective as previous studies: monitoring equipment health and predicting failure times. An integrated physics-based simulation model of machines and sensors is presented (photoelectric, proximity, and vision sensors). Physics-based simulation models based on digital twin concepts can be used to estimate the machine's remaining useful life (RUL).

Modelling, tuning, simulation, and RUL calculation are done in four phases. The accurate operation of the device is used to simulate the model, which is then used to compare the model to the following test setup and to predict any faults that may occur in the future (Aivaliotis et al., 2019).

Since vibration analysis is the most frequently utilized analysis for automatically scheduling Maintenance on manufacturing problems, it is a fundamental PdM activity in the manufacturing industry. Motors and tubes, for example, are attached to vibration sensors, which provide key information about how they are running. In addition, the non-continuous sampling strategy of vibration sensors and the difficulties of interpreting data present novel technical challenges to the analytical system. To maximize the utility and minimize the operational overhead of vibration sensors, a novel analytical framework specially designed for vibration analysis based on its unique characteristics was proposed. The remaining Useful Lifetime (RUL) estimation will be used in data engine to improve replacement scheduling for monitored equipment, a crucial problem in maintaining cyber-physical systems. Vibration data analysis has been shown to extend tube lifetimes by 1.2x and reduce replacement costs by 20% when performed on real manufacturing sites (Jung et al., 2017, pp. 1–5). The following figures describe the learning model and the sensor installation's mechanical setup.

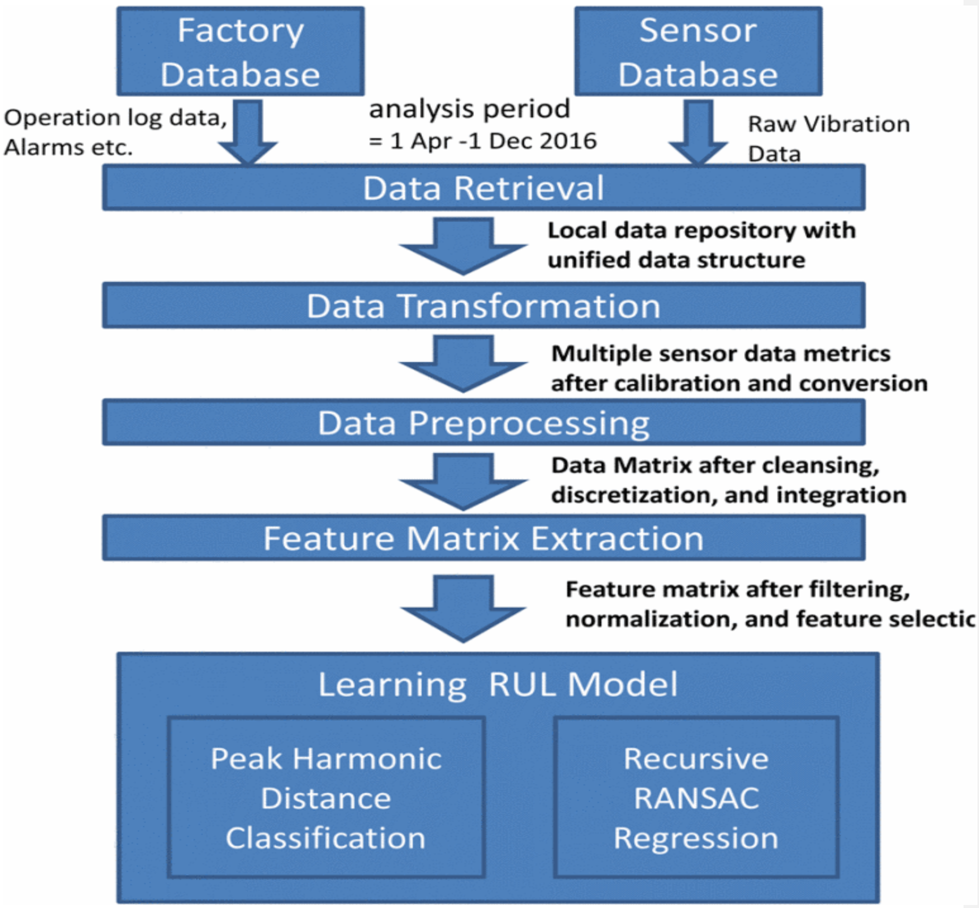

Figure 4-7: Sensor Vibration Analysis Data Acquisition And Processing Chain

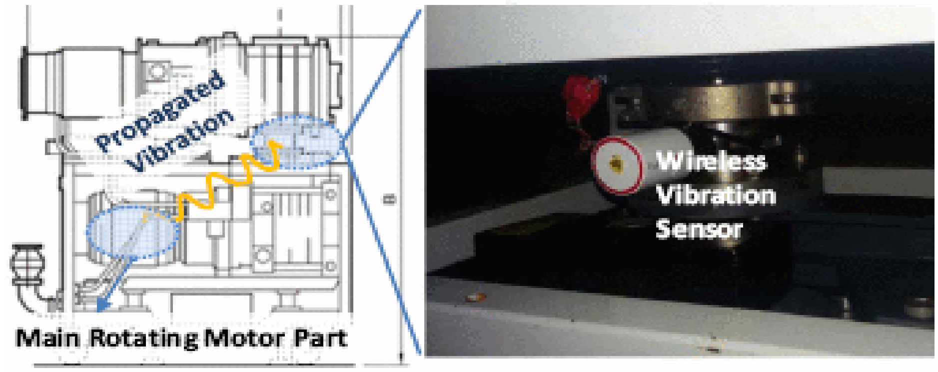

Figure 4-8: Vibration Sensors' Mechanical Orientation

IoT middleware eliminates variations in technological infrastructure by enabling real-time communication and data dissemination between monitored equipment and the system in PdM platforms. This is done by utilizing a variety of communication protocols and interfaces. To continuously gather sensor data from manufacturing equipment in a smart factory, the suggested platform's operational validation is carried out with use cases created explicitly for the purpose. PdM's high-level CPS design architecture is detailed in Figure 4-7.

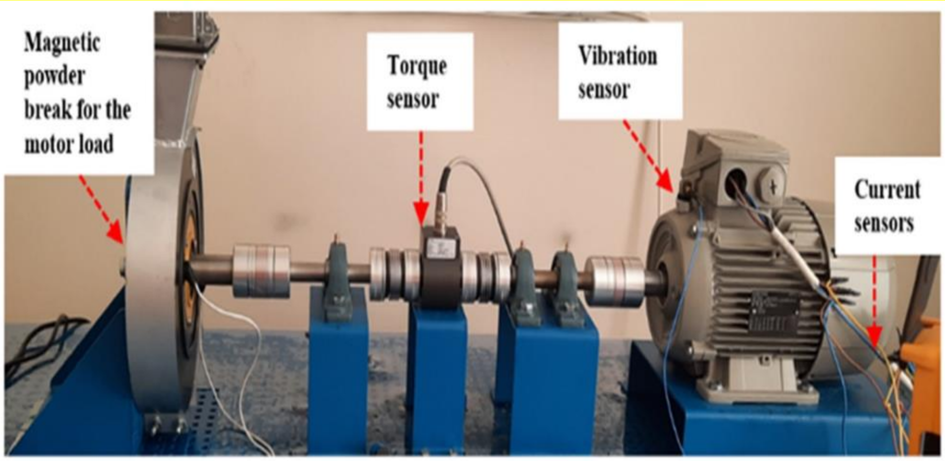

Figure 4-9: Multiple electromechanical systems for motor setup

As an example, MQTT lightweight communication protocol is supported in our electric motors use-case, ROS (Robot Operating System) for any data collection from robotics equipment, and

OPC/UA (Open Platform Communications/Unified Architecture) for industrial automation systems such as Programmable Logic Controllers (PLCs). In an intelligent factory environment involving many processes and equipment fleets, the platform can collect and monitor most of the equipment through these data adapters. The system is designed as a platform where a scalable data distributor component called the context broker acts as middleware with standardized data models for each piece of equipment. Other service components, such as big data persistence software, can be connected to it to receive IoT sensor data (E. Cinar et al., 2022, pp. 1–3).

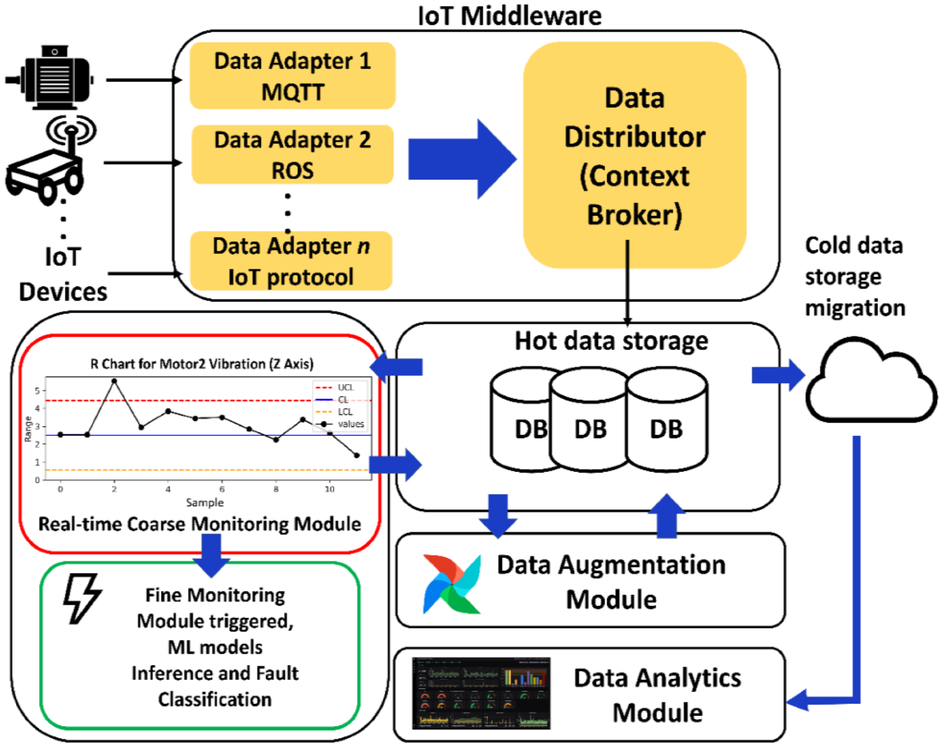

Figure 4-10: Iot PdM Architecture for CPS

In thermal imaging, malfunctioning electrical installations experience a temperature rise. Thermal imaging uses thermal imaging cameras to visually represent the surface temperatures of an object. This representation is helpful for studying the pattern of heat on the surface of an object and for identifying hot spots, i.e. areas that may be problematic. These problem areas can be analyzed and repaired to prevent failures/collapses from causing further damage. As with any technology, thermal imaging has advantages and limitations.

The "Reliability and Maintainability (R&M) Guideline for Manufacturing Machinery and Equipment (M&E)" defines failure as follows in the Society of Automotive Engineers (SAE) guideline: Failure occurs when machinery/equipment and appliances are not able to produce parts under specified conditions or perform scheduled operations when scheduled. Action must be taken after every failure. Two case studies are performed: 1) To soundproof doors, a car assembly plant decided to buy a foam spray device, and 2) A factory is setting up an automatic pallet assembly system with two axes that are rotated by 90 degrees each. Failure mode/ effects, cause, controls and responsibilities are described in detail for both cases, followed by emerging issues in the form of questions and their related suggestions (Blache & Shrivastava, 1994, pp. 69–72). The typical definition of reliability is the possibility that a device will continue functioning as intended for a predetermined period under predetermined conditions. For instance, a probability of 99 percent indicates that, on average, a device will operate correctly 99 out of 100 times. The reliability parameter known as Mean Time Between Failures (MTBF) is well-known. It should be possible to determine what constitutes equipment nonperformance using the term "intended function" used to describe equipment performance. The failure of M&E is the nonperformance that was so agreed upon or perceived. Early in the M&E life cycle, the time frame for the equipment to work dependably should be established. The appliance's operating and environmental circumstances, or stresses, may encounter during its necessary lifetime are referred to as the "performance under stated conditions." The operational conditions must be adequately identified because they differ from one piece of equipment to another. For example, they could be strange voltage transients that enter the equipment through the supply lines. It is vital to define environmental pressures like heat and humidity.

An electromechanical mechanism powers the motor. Since the majority of manufacturing operations are built on motors and engines. It is essential to look for options that lessen the likelihood of failure. These motors are typically quite expensive and have greater efficiency; thus, the industry must function without standby. Motors can be subjected to various predictive techniques to reduce the number of unintended errors. However, the methods based on electrical signature analysis are Motor Current Signature Analysis (MCSA), Expanded Park's Vector Approach (EPVA), and Instant Power Signature Analysis (IPSA), as shown in Figure 4-7. The tests focused on mechanical interpretations are vibration signature analysis, acoustic signature analysis, and speed oscillations signature (Manjare & Patil, 2021, pp. 2–4).

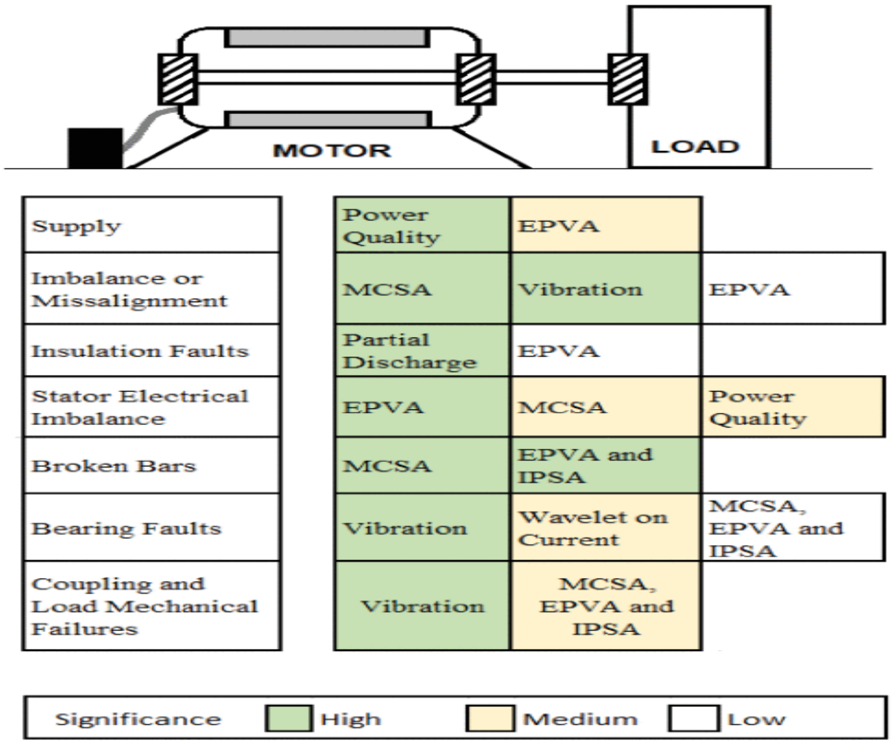

Figure 4-11: Motor vibration signature analysis approaches

Condition monitoring, a crucial component of PdM, can be done manually or with the help of suitable sensors. In the initial step of the methodology's execution, designated as stage 1, phase 1a, sensors are used to gather quantitative data about the equipment that will be monitored. Every vibration frequency, for instance, can be connected to a particular failure since specific frequencies are only present when circumstances point to a potential flaw. Advanced instrumentation, such as vibration analysis tools, or skilled senses, can be used to monitor machinery. Phase 1b, therefore, gathers information from system administrators and specialists to ensure that there is additional knowledge from human details and information. Condition monitoring aims to collect information about an item's state to identify impending failure and schedule repair actions accordingly. Another objective is to increase understanding of failure causes and effects and deterioration patterns. Stage

2 involves analysing and synthesising the information gathered to carry out stages 3 and 4, which have as their goals the establishment of a system description and the determination of objectives and criteria effectiveness. Stage 5 of the testing procedure involves creating a database of the equipment being tested using aggregated and processed data from stage 2. In stage 6, several options are developed through in-depth conversations with industry experts. The results of stages 4, 5, and 6 are used to build a mathematical model of the system under investigation and to develop a suitable optimization job (stage 7).

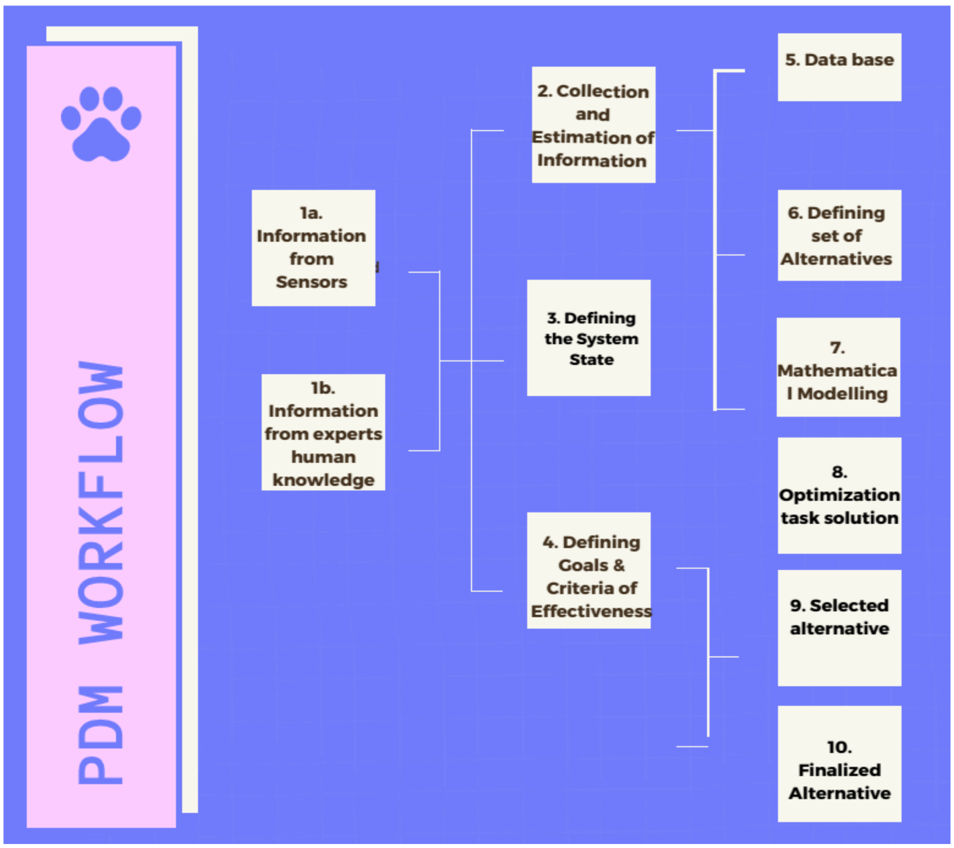

Figure 4-12: PdM workflow

Step eight is the selection of the ideal response. At stage 9, the defined solution alternative is applied. In stage 10, the alternative application is contrasted with the goals. The optimization procedure for decision-making is restarted at step 7, and steps 8, 9, and 10 are repeated if the outcomes are different from what was anticipated and the relevant correction is placed in the mathematical model of the system. In some instances, redefining the objectives and criteria in stage 4 is necessary to meet the desired goals for the management of the system because the a priori information about the researched system is insufficient. The structural diagram in Fig. 2 helps illustrate how stages 6, 7, 8, 9, and 10 are implemented. A group or a single option is chosen and assessed against all the objectives resulting from optimization solutions. Finally, the expert must approve the preferred choice or resort (designated alternative).

5. Results and Discussions

PdM is skillfully becoming an essential ingredient for industry 4.0 implementation. It shapes many conventional industries, such as manufacturing, oil, elevators, electrical and mechanical industries, and plants. The rapid development of LoT technology and the rise of deep learning have greatly expanded their applications in various industries.

Different maintenance types are described, with Run Two Failure (R2F), Preventive Maintenance (PvM), Condition Based Maintenance (CbM), and PdM serving as the main categories.

The outcomes demonstrate that installing PdM based on deep learning is very effective for automatic PdM systems and enhances the plant's or manufacturing system's system efficiency. The vibration measurements, for instance, can reveal imbalance, misalignment, and broken components on bearings when used in vibration analysis.

Two real-world use cases, including monitoring autonomous transfer trucks and electric motor equipment for an intelligent industrial environment, are tested with the deep learning-based CPS implementation of PdM. KPIs and system metrics are defined to illuminate further the end-users performance and maintenance requirements in CPSs. The recommended remedy demonstrates the integration of sophisticated problem-detection systems for intelligent production settings using CPSs and PdMsystems. The extraordinarily general and adaptable CPS design allows the proposed PdM system to accommodate different software stacks easily.

The research shed light on the major deep learning technologies and IoT implementation for future PdM system successful implementation for Industry 4.0. The methods studied include SVM, Random Forest, Recurrent Neural Networks, and LSTM. Considering the importance of big data for PdM implementation, the potential of big data for PdM implementation has been successfully explored. The challenges in the successful implementation of PdM in different sectors are reviewed. Later, the different types of analysis in the PdM for manufacturing, oil plants and mechanical sectors are reviewed, such as vibration analysis and oil analysis as the main analysis types. Their limitation and theoretical background are evaluated in a futuristic research-oriented approach. Infra-red thermography analysis, and current and ultrasonic analysis are other types studied in the research. Finally, the study concluded with the failure mode analysis (or Failure Mode Effect Analysis) section and performed three different case studies as a benchmark for successful PdM implementation research.

The study kicked off with the maintenance philosophy and the maintenance evolution in the context of industry 4.0. The maintenance types are reviewed, and PdM is emphasized as a viable solution in the implementation effort in the context of industry 4.0; the applications of machine learning and deep learning methods are surveyed in this research. The algorithms such as SVM, Recurrent neural networks (LSTM and Gated Neural Networks), and Convolutional Neural Networks have practical applications in the next generation of maintenance.

The fault diagnosis, prognosis, and failure detection are the mechanism that is well suited for a machine learning technique called anomaly detection. If necessary, this technique exploits the thresholding concept and triggers an alarm in case of irregular patterns. The model learns these patterns of the PdM system through training on large datasets obtained from sensors. This brings the question of errors, prediction accuracy, and noisy and unlabeled raw, unlabeled data that needs pre-processing steps before being fed to the failure prediction models. The debate started with the question of which models and techniques are best suited for which type of application in industry, ranging from oil, manufacturing energy, mechanical and electrical. This debate was clarified with extensive research using keyword search bases techniques on popular and bigger intellectual and scientific papers databases and tables, and data were extracted. Case studies are performed to resolve the best possible algorithm for a particular PdM application in the industry.

6. Conclusion & Future Recommendations

In the last twenty years, the shift from preventive maintenance towards condition-based monitoring has largely been aided by technological advancement (e.g. deep learning and IoT) and shifts towards digitization. Many technologies contribute to this shift in production, plant and manufacturing maintenance, including CPS (Cyber-Physical systems), the Industrial

Internet of Things, big data analytics, cloud computing, fog computing, edge computing, machine learning, artificial intelligence, digital twins, and neural-inspired methodologies.

The PdM mindset has primarily been adopted by energy, wind turbine plants, oil plants, and manufacturers mainly due to possible collaborations and AI solutions provided by AI solutions development companies.

In the future, mechanical, electrical, energy, and cement manufacturers will deploy PdMstrategies on a large scale and possibly even develop divisions within the company. By leveraging their data analytics and processing skills, AI creators can work through large amounts of data and extract results from it.

Under the Umbrella of Industry 4.0, PdM will eventually gracefully integrate all of these features and their subtleties into a single product that proactively and anticipatorily manages asset maintenance, enhances uptime, and optimizes productivity in an increasingly connected manufacturing ecosystem.

PdM was worth just 1.5 billion globally five years ago. Most systems operated independently and were made up of 'Do It Yourself' type data-science solutions, limiting the advantages of the technology due to a lack of compatibility and comprehension.

The PdMindustry is expected to overgrow over the next several years as it focuses on extending the life of failing industrial machinery. In response to the increase in demand, asset uptime must be maximized while maintenance costs are reduced. The rules, regulations, and standards that must be followed in industrial manufacturing plants have become more severe. As a result, PdM in the workplace is an effective tool.

- Industry 4.0, the loT, and artificial intelligence are expected to help the industry and manufacturing sector's revenue proliferate, reaching a value of $28.2 billion by 2026. It is also possible to eventually be integrated into enterprise software, or businesses will choose best-of-breed solutions that work seamlessly with current systems. According to experts' projections, corporate asset management products will include 60 percent of IoT-enabled PDM solutions by 2026, up from 15 percent in 2021.
- Globally, manufacturing and other energy-intensive industries are crucial to tackling climate change and advancing sustainability.

- Advances in the three critical areas of sustainability—reducing the amount of materials and spares wasted, cutting down on the energy used in manufacturing and industrial processes, extending the operational lifetime of assets, and taking appropriate steps to lower health and safety risks—will be made possible by real-time data, analytics, and PdM technologies..

- PdM programs will become as essential to industrial organizations as enterprise resource planning (ERP) or financial planning software. This translates into equipment performance that demonstrates best practices, adheres to industry standards, and generates a competitive advantage.

- PdM and Condition Monitoring (CM) are the industrial ecosystem's most researched, implemented, and top opportunities. PDM & CM digitization applies advanced computing technology to improve machine efficiency, reliability, outcomes and sustainability. Together with its components, it encompasses a collaborative approach to a man-machine-technology ecosystem.

6.1 PdM Advantages and Implementation for Future Systems

PdM does not require a complete restructuring of your company's infrastructure. Instead, it may be added to your already-existing infrastructure for integrated control and information.

The first stage in the process is to debate and decide what data you want to collect. The historical data needed for this can be obtained through sensors, industrial machinery, defect records, collapses, failures, or other difficulties you wish to predict, as well as failures that have already happened. The necessary historical data is made available via sensor readings, industrial equipment, and defect records.

PdM analytics software identifies the fundamental cause and early warning symptoms based on data from past downtime issues. To conclude, analytics software monitors data traffic via "agents" deployed locally in the plant or the cloud.

There are two types of agents used by analytics software. The first kind is failure agents, which keep an eye out for patterns that can foretell a failure in the future. If such behaviours are detected, the agent will notify or message plant staff and administer the recommended remedy.

Anomaly agents, which keep an eye on regular operating procedures and seek changes, including adjustments to operational or environmental parameters, are the second category. These agents also inform staff members of any observed changes so that they can look into them and, if required, take corrective action.

6.2 Using Crystal Ball/Horoscope/Predictive Capability For Foreseeing The Future Of PdM

Years have passed since the invention of predictive technology. In addition to helping us do online searches, it is used to identify credit card fraud and optimise marketing campaigns. Its purpose in the industrial sector is to meticulously record incidents and failures so that they can be recognised and fix equipment or machine problems as soon as they arise.

Historical failure reports have great potential for reducing future failures and downtime, as many manufacturers already recognise. Utilizing this information, which is already there in your assets, can also help you cut down on needless maintenance and operations risks while lowering unexpected downtime.

REFERENCES

Achouch, M., Dimitrova, M., Ziane, K., Sattarpanah Karganroudi, S., Dhouib, R., Ibrahim, H., & Adda, M. (2022). On PdM in Industry 4.0: Overview, Models, and Challenges. Applied Sciences, 12(16), 8081. https://doi.org/10.3390/app12168081.

Kane, A. P., Kore, A. S., Khandale, A. N., Nigade, S. S., & Joshi, P. P. (2022). PdMusing Machine Learning. *arXiv preprint arXiv:2205.09402*.

Cheng, X., Chaw, J. K., Goh, K. M., Ting, T. T., Sahrani, S., Ahmad, M. N., Abdul Kadir, R., & Ang, M. C. (2022). Systematic Literature Review on Visual Analytics of PdM in the Manufacturing Industry. Sensors (Basel, Switzerland), 22(17), 6321. https://doi.org/10.3390/s22176321.

Kubiak, K., Dec, G., & Stadnicka, D. (2022). Possible Applications of Edge Computing in the Manufacturing Industry-Systematic Literature Review. Sensors (Basel, Switzerland), 22(7), 2445. https://doi.org/10.3390/s22072445

A, S. (2020). Review on Industrial Internet of Things (IIOT). *International Journal for Research in Applied Science and Engineering Technology*, *8*, 523–525. https://doi.org/10.22214/ijraset.2020.2080

Abidi, M. H., Mohammed, M. K., & Alkhalefah, H. (2022). Predictive Maintenance Planning for Industry 4.0 Using Machine Learning for Sustainable Manufacturing. *Sustainability*, *14*(6). https://doi.org/10.3390/su14063387

Albon, C. ,. (2018). Support Vector Machines. In *Machine Learning with Python Cookbook* (pp. 282–293). Oreilly.

Awatramani, J., Verma, G., Hasteer, N., & Sindhwani, R. (2022). Investigating Strategies and Parameters to Predict Maintenance of an Elevator System. In V. V. Rao, A. Kumaraswamy, S. Kalra, & A. Saxena (Eds.), *Computational and Experimental Methods in Mechanical Engineering* (pp. 323–332). Springer Singapore.

Bevilacqua, M., Braglia, M., & Gabbrielli, R. (2000). Monte Carlo simulation approach for a modified FMECA in a power plant. *Quality and Reliability Engineering International*, *16*(4), 313–324.

Blache, K. M., & Shrivastava, A. B. (1994). Defining failure of manufacturing machinery and equipment. *Proceedings of Annual Reliability and Maintainability Symposium (RAMS)*, 69–75. https://doi.org/10.1109/RAMS.1994.291084

Boldosova, V., Hietala, J., Pakkala, J., Salokangas, R., Kaarmila, P., & Puranen, E. (2021). Predictive Analytics in the Production of Elevators. In T. Cerquitelli, N. Nikolakis, N. O'Mahony, E. Macii, M. Ippolito, & S. Makris (Eds.), *Predictive Maintenance in Smart Factories: Architectures, Methodologies, and Use-cases* (pp. 165–185). Springer Singapore. https://doi.org/10.1007/978-981-16-2940-2_8

Bousdekis, A., Lepenioti, K., Apostolou, D., & Mentzas, G. (2019). Decision Making in Predictive Maintenance: Literature Review and Research Agenda for Industry 4.0. *IFAC-PapersOnLine*, *52*(13), 607–612. https://doi.org/https://doi.org/10.1016/j.ifacol.2019.11.226

Bowles, J. B. (1998). The new SAE FMECA standard. *Annual Reliability and Maintainability Symposium. 1998 Proceedings. International Symposium on Product Quality and Integrity*, 48–53. https://doi.org/10.1109/RAMS.1998.653561

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Cachada', 'A. (2018). "Intelligent and predictive maintenance in manufacturing systems." *Nan*, *nan*.

Canizo, M., Onieva, E., Conde, A., Charramendieta, S., & Trujillo, S. (2017). Real-time predictive maintenance for wind turbines using Big Data frameworks. *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*, 70–77. https://doi.org/10.1109/ICPHM.2017.7998308

Chukwuekwe, D. O. (2016). *Condition Monitoring for Predictive Maintenance: - A Tool for Systems Prognosis within the Industrial Internet Applications*.

Chuprina R, K. O. (2021, July 23). *The Complete Guide to Predictive Maintenance with Machine Learning*. https://spd.group/machine-learning/predictive-maintenance/

Cinar, E., Kalay, S., & Saricicek, I. (2022). A Predictive Maintenance System Design and Implementation for Intelligent Manufacturing. *Machines*, *10*(11). https://doi.org/10.3390/machines10111006

Cinar, Z., Nuhu, A., Zeeshan, Q., Korhan, O., Asmael, M., & Safaei, B. (2020). Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0. *Sustainability*, *12*, 8211. https://doi.org/10.3390/su12198211

de Queiroz Souza, R., & Alvares, A. J. (2007). *FMEA AND FTA ANALYSIS FOR APPLICATION OF THE RELIABILITY-CENTERED MAINTENANCE METHODOLOGY : CASE STUDY ON HYDRAULIC TURBINES*.

Domingues, R., Filippone, M., Michiardi, P., & Zouaoui, J. (2018). A comparative evaluation of outlier detection algorithms: Experiments and analyses. *Pattern Recognition*, *74*, 406–421.

Ellis, B. A., & Byron, A. (2008). Condition based maintenance. *The Jethro Project*, *10*, 1–5.

Géron, A. ,. (2019). Unsupervised Learning. In *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow* (Second, pp. 238–274). Oreilly Media Inc.

Gilabert, E., & Arnaiz, A. (2006). Intelligent automation systems for predictive maintenance: A case study. *Robotics and Computer-Integrated Manufacturing*, *22*(5), 543–549. https://doi.org/https://doi.org/10.1016/j.rcim.2005.12.010

Giuliano Liguori. (2022, March 23). Predictive maintenance in industry 4.0: applications and advantages. *Linkedin.Com*.

Gonfalonieri, A. ,. (2019, November 7). *How to Implement Machine Learning For Predictive Maintenance*. https://towardsdatascience.com/how-to-implement-machine-learning-for-predictive-maintenance-4633cdbe4860

Hashemian, H. M. (2011). State-of-the-Art Predictive Maintenance Techniques. *IEEE Transactions on Instrumentation and Measurement*, *60*(1), 226–236. https://doi.org/10.1109/TIM.2010.2047662

Hrnjica, B., & Softic, S. (2020). *Explainable AI in Manufacturing: A Predictive Maintenance Case Study* (pp. 66–73). https://doi.org/10.1007/978-3-030-57997-5_8

Hung, Y. H. (2021). Improved ensemble-learning algorithm for predictive maintenance in the manufacturing process. *Applied Sciences (Switzerland)*, *11*(15). https://doi.org/10.3390/app11156832

J. Levitt. (2011). *Complete Guide to Preventive and Predictive Maintenance* (J. Carleo, J. Romano, & R. Weinstein, Eds.; 2nd ed.). Industrial Press Inc.

Julianna D. (2021). *Supervised vs. Unsupervised Learning: What's the Difference?* https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning

Jung, D., Zhang, Z., & Winslett, M. (2017). Vibration Analysis for IoT Enabled Predictive Maintenance. *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, 1271–1282. https://doi.org/10.1109/ICDE.2017.170

Jürgensen, J., & Weigold, M. (2021). A Systematic Review on Predicting and Forecasting the Electrical Energy Consumption in the Manufacturing Industry. *Energies*, *14*, 968. https://doi.org/10.3390/en14040968

Kalligeros, S. (2013). Predictive Maintenance of Hydraulic Lifts through Lubricating Oil Analysis. *Machines*, *2*, 1–12. https://doi.org/10.3390/machines2010001

Kennedy, S. ,. (2020, October 21). *6 case studies illuminate the value of predictive and prescriptive maintenance*. https://www.plantservices.com/predictive-maintenance/predictive-maintenance/article/11290707/6-case-studies-illuminate-the-value-of-predictive-and-prescriptive-maintenance

Kovaříková, I., Szewczyková, B., Blaškoviš, P., Hodúlová, E., & Lechovič, E. (2009). Study and characteristic of abrasive wear mechanisms. *Materials Science and Technology*, *1*, 1–8.

Küfner, T., Döpper, F., Müller, D., & Trenz, A. (2021). Predictive Maintenance: Using Recurrent Neural Networks for Wear Prognosis in Current Signatures of Production Plants. *International Journal of Mechanical Engineering and Robotics Research*, 583–591. https://doi.org/10.18178/ijmerr.10.11.583-591

Lai, C. T. A., Jiang, W., & Jackson, P. R. (2019). Internet of Things enabling condition-based maintenance in elevators service. *Journal of Quality in Maintenance Engineering*, *25*(4), 563–588. https://doi.org/10.1108/JQME-06-2018-0049

Lee, C., Cao, Y., & Ng, K. K. H. (2017). *Big Data Analytics for Predictive Maintenance Strategies*. https://doi.org/10.4018/978-1-5225-0956-1.ch004

Lisowski, E. (2022, April 14). *Predictive Maintenance and Prevention using Machine Learning (update: 2022)*. Addepto.Com. https://addepto.com/blog/predictive-maintenance-prevention-machine-learning/

Lughofer, E., & Sayed Mouchaweh, M. (2019). Predictive Maintenance in Dynamic Systems Advanced Methods, Decision Support Tools and Real-World Applications: Advanced Methods, Decision Support Tools and Real-World Applications. In *Predictive Maintenance in Dynamic Systems: Advanced Methods, Decision Support Tools and Real-World Applications*. https://doi.org/10.1007/978-3-030-05645-2

Luthra, P. (1991). FMECA: an integrated approach. *Annual Reliability and Maintainability Symposium. 1991 Proceedings*, 235–241.

Ma, X., Chengkai, L., Ng, K. H., & Tan, H.-P. (2021). An Internet of Things-Based Lift Predictive Maintenance System. *IEEE Potentials*, *40*(1), 17–23. https://doi.org/10.1109/MPOT.2020.2973697

Malawade, A. V., Costa, N. D., Muthirayan, D., Khargonekar, P. P., & al Faruque, M. A. (2021). Neuroscience-Inspired Algorithms for the Predictive Maintenance of Manufacturing Systems. *IEEE Transactions on Industrial Informatics*. https://doi.org/10.1109/TII.2021.3062030

Manjare, A. A., & Patil, B. G. (2021). A Review: Condition Based Techniques and Predictive Maintenance for Motor. *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, 807–813. https://doi.org/10.1109/ICAIS50930.2021.9395903

Mathworks. (2021). *Overcoming Four Common Obstacles to Predictive Maintenance with MATLAB and Simulink*. https://www.mathworks.com/content/dam/mathworks/white-paper/gated/predictive-maintenance-challenges-whitepaper.pdf

Misra, K. B., & Weber, G. G. (1989). A new method for fuzzy fault tree analysis. *Microelectronics Reliability*, *29*(2), 195–216. https://doi.org/https://doi.org/10.1016/0026-2714(89)90568-4

Mitul M. (2020, February 13). *How to Find the Right Machine Learning Techniques for Predictive Maintenance?* AI TECHNOLOGY INSIGHTS. https://aithority.com/guest-authors/how-to-find-the-right-machine-learning-techniques-for-predictive-maintenance/

Mobley, R. K. (2002). 6 - Predictive Maintenance Techniques. In R. K. Mobley (Ed.), *An Introduction to Predictive Maintenance (Second Edition)* (pp. 99–113). Butterworth-Heinemann. https://doi.org/https://doi.org/10.1016/B978-075067531-4/50006-3

Nakayama, K., & Martin, J.-M. (2006). Tribochemical reactions at and in the vicinity of a sliding contact. *Wear*, *261*(3), 235–240. https://doi.org/https://doi.org/10.1016/j.wear.2005.10.012

Namuduri, S., Narayanan, B. N., Davuluru, V. S. P., Burton, L., & Bhansali, S. (2020). Review—Deep Learning Methods for Sensor Based Predictive Maintenance and Future Perspectives for Electrochemical Sensors. *Journal of The Electrochemical Society*, *167*(3), 037552. https://doi.org/10.1149/1945-7111/ab67a8

Nangia, S., Makkar, S., & Hassan, R. (2020). IoT based Predictive Maintenance in Manufacturing Sector. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3563559

Nguyen, K. T. P., & Medjaher, K. (2019). A new dynamic predictive maintenance framework using deep learning for failure prognostics. *Reliability Engineering & System Safety*, *188*, 251–262. https://doi.org/https://doi.org/10.1016/j.ress.2019.03.018

OSA-CBM. (2022). *Open System Architecture for Condition-Based Maintenance*. MIMOSA. https://www.mimosa.org/mimosa-osa-cbm/

Prihatno, A. T., Nurcahyanto, H., & Jang, Y. M. (2021). Predictive Maintenance of Relative Humidity Using Random Forest Method. *2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 497–499. https://doi.org/10.1109/ICAIIC51459.2021.9415213

Rieger, T., Regier, S., Stengel, I., & Clarke, N. (2019). Fast Predictive Maintenance in Industrial Internet of Things (IIoT) with Deep Learning (DL): A Review. *CERC*.

Rivas, A., Fraile, J., Chamoso, P., González Briones, A., Sittón-Candanedo, I., & Corchado Rodríguez, J. (2020). *A Predictive Maintenance Model Using Recurrent Neural Networks* (pp. 261–270). https://doi.org/10.1007/978-3-030-20055-8_25

Sang, G. M., Xu, L., & de Vrieze, P. (2021). A Predictive Maintenance Model for Flexible Manufacturing in the Context of Industry 4.0. *Frontiers in Big Data*, *4*. https://doi.org/10.3389/fdata.2021.663466

Serradilla, O., Zugasti, E., Rodriguez, J., & Zurutuza, U. (2022). Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects. *Applied Intelligence*, *52*(10), 10934–10964. https://doi.org/10.1007/s10489-021-03004-y

Sharma, R., & Sharma, P. (2010). System failure behavior and maintenance decision making using, RCA, FMEA and FM. *Journal of Quality in Maintenance Engineering*, *16*, 64–88. https://doi.org/10.1108/13552511011030336

Shen, L. J., Lukose, J., & Young, L. C. (2021). Predictive maintenance on an elevator system using machine learning. *Journal of Applied Technology and Innovation (e-ISSN: 2600-7304)*, *5*(1), 75.

Spreafico, C., Russo, D., & Rizzi, C. (2017). A state-of-the-art review of FMEA/FMECA including patents. *Computer Science Review*, *25*, 19–28. https://doi.org/https://doi.org/10.1016/j.cosrev.2017.05.002

Spronk, D. (2022). *Predictive Maintenance*. Deloitte. https://www2.deloitte.com/cz/en/pages/deloitte-analytics/solutions/predictive-maintenance.html

Sundeep R. (2022, February 22). *Unplanned Downtime Costs More Than You Think*. https://www.forbes.com/sites/forbestechcouncil/2022/02/22/unplanned-downtime-costs-more-than-you-think/?sh=1bb32aab36f7

Sutrisno, A., Gunawan, I., & Tangkuman, S. (2015). Modified Failure Mode and Effect Analysis (FMEA) Model for Accessing the Risk of Maintenance Waste. *Procedia Manufacturing*, *4*, 23–29. https://doi.org/https://doi.org/10.1016/j.promfg.2015.11.010

Taylor, K. ,. (2022, November 11). *What is the Importance of Predictive Maintenance in Industry 4.0?*

Thakker, D., Patel, P., Intizar Ali, M., Shah, T., Cao, Q., Samet, A., Zanni-Merk, C., de Bertrand De Beuvron, F., & Reich, C. (2020). Combining chronicle mining and semantics for predictive maintenance in manufacturing processes. *Semantic Web*, *11*(6). https://doi.org/10.3233/SW-200406

Tshabuse, F., & Pretorius, J. H. C. (2013). *Applying preventive and predictive best practice on plant maintenance*.

vroc.ai team. (2022). *How Predictive Maintenance is driving Industry 4.0*. Vroc.Ai. https://vroc.ai/how-predictive-maintenance-is-driving-industry-4-0/

Wang, J., Zhang, L., Duan, L., & Gao, R. X. (2017). A new paradigm of cloud-based predictive maintenance for intelligent manufacturing. *Journal of Intelligent Manufacturing*, *28*(5). https://doi.org/10.1007/s10845-015-1066-0

Wang, K., Dai, G., & Guo, L. (2016). Intelligent Predictive Maintenance (IPdM) for Elevator Service- Through CPS, IOT&S and Data Mining. *Proceedings of the 6th International Workshop of Advanced Manufacturing and Automation*, 1–6. https://doi.org/https://doi.org/10.2991/iwama-16.2016.1

Wang, K., & Wang, Y. (2018). How AI Affects the Future Predictive Maintenance: A Primer of Deep Learning. In K. Wang, Y. Wang, J. O. Strandhagen, & T. Yu (Eds.), *Advanced Manufacturing and Automation VII* (pp. 1–9). Springer Singapore.

Wang, W., Scarf, P. A., & Smith, M. A. J. (2000). On the application of a model of condition-based maintenance. *Journal of the Operational Research Society*, *51*(11), 1218–1227. https://doi.org/10.1057/palgrave.jors.2601042

Zhang, S., Zhou, J., Wang, E., Zhang, H., Gu, M., & Pirttikangas, S. (2022). State of the art on vibration signal processing towards data-driven gear fault diagnosis. *IET Collaborative Intelligent Manufacturing*, *n/a*(n/a). https://doi.org/https://doi.org/10.1049/cim2.12064

Zhang, W., Yang, D., & Wang, H. (2019). Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey. *IEEE Systems Journal*, *13*(3), 2213–2227. https://doi.org/10.1109/JSYST.2019.2905565