EVOLUTIONARY ALGORITHMS FOR IMPROVING ENERGY EFFICIENCY AND SECURITY IN WIRELESS SENSOR NETWORKS

A Thesis Submitted to

Delhi Technological University

for the award of the degree of

Doctor of Philosophy

In

Electronics and Communication Engineering

By

RAJIV YADAV

(Enrollment No.: 2K17/PHDEC/04)

Under the supervision of

DR. INDU SREEDEVI
Professor and Dean (Student Welfare)

DR. DAYA GUPTA
Professor in Dept. of CSE

Department of Electronics & Communication Engineering

Delhi Technological University

Delhi-110042, India

AUGUST 2023

Delhi Technological University

Formerly Delhi College of Engineering

Shahbad Daulatpur, Bawana Road, Delhi –110042

CERTIFICATE

This is to certify that the thesis entitled "**Evolutionary Algorithms for Improving Energy Efficiency and Security in Wireless Sensor Networks"**, being submitted by **Rajiv Yadav** (Reg. No.: 2K17/PHDEC/04) to the Department of Electronics & Communication Engineering, Delhi Technological University in partial fulfilment of the requirements for the award of the degree of **Doctor of Philosophy** is based on the original research work carried out by him under the guidance and supervision of **Prof. S. Indu** and **Prof. Daya Gupta.** In our opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree. It is further certified that the work embodied in this thesis has not been submitted, in part or full, to any other university or institution for the award of any degree or diploma.

Prof. Indu Sreedevi
Supervisor
Department of ECE
Delhi Technological University,
Delhi –110042, India

Prof. Daya Gupta
Co-Supervisor
Department of CSE
Delhi Technological University,
Delhi –110042, India

Prof. O.P. Verma
Head of the Department
Department of ECE
Delhi Technological University
Delhi –110042, India

Delhi Technological University

Formerly Delhi College of Engineering

Shahbad Daulatpur, Bawana Road, Delhi –110042

DECLARATION OF AUTHORSHIP

I hereby declare that all information in the thesis entitled "**Evolutionary Algorithms for Improving Energy Efficiency and Security in Wireless Sensor Networks**" has been obtained and presented in accordance with the academic rules and ethical conduct laid out by Delhi Technological University. I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Rajiv Yadav

2K17/PHDEC/04

Department of Electronics & Communication Engineering

Delhi Technological University

Delhi –110042

ACKNOWLEDGEMENTS

First and foremost, thanks to the Almighty for giving me strength and inspiration to carry out this research work. I owe a deep sense of gratitude to all his comprehensive soul whose divine light has enlightened my path throughout the journey of my research.

I take the opportunity to humbly submit my sincere and heartfelt thanks to my research supervisors **Prof. S. Indu** & **Prof. Daya Gupta** from the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, and the Department of Computer Science Engineering, Delhi Technological University, Delhi for their valuable guidance, enthusiastic encouragement, and persistent support. I am truly grateful from the core of my heart for their meticulous approach, wonderful assistance of their perspective, and fruitful discussions on the research topic. Their immense contribution and rare dedication in providing the much-needed guidance, is worth of much honour. Their careful supervision and personal attention have given me a lot of confidence and enthusiasm, during the different stages of my doctoral investigations. They are academic giants under whose watch am moulded to a seasoned research scholar. I invariably fall short of words to express my sincere gratitude for their patience and motivation.

I am extremely thankful to **Prof. O. P. Verma, Head of the Department of Electronics and Communication Engineering,** Delhi Technological University, Delhi, and other faculty members for their endless support and cooperation throughout this dissertation. I express my sincere gratitude to **Ms. Ishu Tomar** and **Ms. Palak Handa,** Department of ECE, Delhi Technological University, Delhi, for constantly encouraging and supporting me in this endeavour of mine. I also thank the management of Delhi Technological University, Delhi, for countenancing me to pursue Ph.D. at the Department of ECE, Delhi Technological University, Delhi.

I am greatly indebted to my dear friends **Mr. Kiros** and **Mr. Amarendra**, who constantly supported, counselled, and encouraged me all the time throughout this research period. My appreciation also goes to my colleagues in the Dept. of Electronics and Communication Engineering, Delhi Technological University, Delhi, for their constant guidance and assistance especially Amit Gautam, Barkha Singh, Kavita Bhatt, and Shashank. I am thankful to all staff members of the department of ECE for their kind help and support during the entire period of my research.

I dedicate this thesis to my family for their endless love, support, encouragement, and blessings throughout my academics. My mother, **Smt. Sushila Devi** without whom I could not imagine being enrolled in Ph.D., for her unwavering encouragement, undeterred faith in me, and being the biggest pillar of strength who supported me all the way till the end. My friend, **Ms. Neeraj Yadav** who always wanted me to be a doctor has been a source of motivation and strength during moments of despair and discouragement. I am indeed grateful to my father **Shri Birender Singh**, and my brother **Mr. Sanjeev** for their continuous support, care, and warm wishes for my life. I feel very lucky to have a great family who has supported me throughout the journey of my research work.

Last but not least, I acknowledge each one of those who directly or indirectly have helped me during the whole period, thus making it a well-rounded experience of learning. Many thanks to all of you, this thesis would have no sense without any person I named!

Rajiv Yadav

This Thesis is Dedicated to My Mother.

For her endless love, support and encouragement

ABSTRACT

Wireless Sensor Network (WSN) finds vast real-world applications in the field of energy control, security, health care, defense, and environment monitoring. WSNs are subdued by limited power with a specific battery backup. Due to the large distance between Sensor Nodes (SNs) and the sink, more power consumption occurs in the sensors. The limited energy of SNs is a major drawback to empower a large network coverage area. Therefore, the battery life and location of Cluster Heads (CHs) play an important role in increasing the efficiency and lifetime of SNs for long-term operation in WSNs. Researchers face significant challenges in developing more energy-efficient and secure clustering and routing protocols for WSNs. The bulk of existing routing protocols focuses on CH election while disregarding other important aspects of routing including cluster formation, data aggregation, and security, among others. Nature-inspired algorithms like Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Gray Wolf Optimization (GWO), and Butterfly Optimization Algorithm (BOA) have been used for addressing important challenges in WSNs, such as sensor lifespan, transmission distance, and energy consumption. The introduction of hybrid techniques has proven to be more effective. Hybrid techniques like GA-PSO, PSO-ACO, PSO-GWO, etc. have gained traction as viable solutions for bio-inspired algorithms to improve the energy efficiency of WSNs. Such techniques have been found more optimum in comparison to the conventional techniques. The present thesis focuses on comparative analysis of hybridization techniques with conventional techniques for improving the energy efficiency of WSNs. In addition, an Improved Butterfly Optimization Algorithm (IBOA) has been proposed for global optimization problems in WSNs. Lastly, the present thesis focuses on addressing threats and security issues in WSNs through a feature selection and Machine Learning (ML) based intrusion detection pipeline. The Fast Correlation-based Feature Selection (FCBFS) has been utilized as the feature selection method. The ML classifiers include Decision Tree (DT), Random Forest (RF),

Naïve Bays (NB), Extra Tree (ET), and Extreme Gradient Boosting (XG-Boost). The contributions presented in this thesis are outlined below:

- We propose a comprehensive review of Bio-inspired Hybrid Optimization Algorithms for Energy-Efficient WSN. We have aimed to discuss and compare various newly implemented, conventional, and hybrid methodologies for establishing a robust energy-efficient WSN wherein parameters like packet loss, energy, throughput, delay, and overhead have been utilized. Various open issues and challenges in WSN development using bio-inspired optimization techniques such as network stability, network dynamic character, secure transmission lines, methods to improve QoS, etc., have been addressed.
- We propose a variable sensor modality IBOA for global optimization problems. The modified optimization approach focuses on unconstrained issues, performs on restricted problems, and remains the future scope of this work.
- We propose an FCBFS method with XG-Boost for the National Security Laboratory-Knowledge Discovery Dataset (NSL-KDD) intrusion detection benchmark dataset to address the threats and security issues in a complicated WSN for IoT applications. Evaluation metrics such as accuracy, precision, recall, and F1-Score have been calculated to gauge the performance and robustness of the proposed research work. A classic accuracy score of 99.84% is achieved in the case of the XG-Boost classifier, wherein the best ten obtained features were selected after applying the proposed FCBFS. The proposed technique, which has ten features, outperforms the existing techniques in the literature for the NSL-KDD dataset. Experimental analysis has been done extensively to prove the efficacy of the developed solutions.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

WSNs	Wireless Sensor Networks
SNs	Sensor Nodes
BS	Base Station
CH	Cluster Head
IoT	Internet of Things
QoS	Quality of Service
DE	Differential Evolution
GA	Genetic Algorithm
PSO	Particle Swarm Optimization
ACO	Ant Colony Optimization
BBO	Biogeography-Based Optimization
GWO	Grey Wolf Optimization
BOA	Butterfly Optimization Algorithm
FS	Feature Selection
FCBFS	Fast Correlation-Based Feature Selection
SEP	Stable Election Protocol
EECS	Energy Efficient Clustering Scheme
LEACH-SWDN	Low Energy Adaptive Clustering Hierarchy with Sliding Window and Dynamic Nodes
DT	Decision Tree
RF	Random Forest

NB	Naïve Bays
ET	Extra Tree
XG-Boost	Extreme Gradient Boosting
ABC	Artificial Bee Colony
HSA	Harmony Search Algorithm
FPA	Flower Pollination Algorithm
HGWSFO	Hybrid GWO-based Sunflower Optimization
HIDS	Hybrid Intrusion Detection System
SVM	Support Vector Machine
DNN	Deep Neural Network
RBM	Restricted Boltzmann Machine
LR	Logistic Regression
ML	Machine Learning
DDoS	Distributed Denial of Service
ETD	Effective Transmission Distance
GABEEC	GA Based Energy Efficient Clustering
PSOBS	PSO Based Selection
PSO-ECSM	PSO-based Energy Efficient Clustering and Sink Mobility
GLBCA	Gateway Load Balanced Clustering Protocol
PSO-UFC	PSO-based Unequal Fault-tolerant Clustering
LDC	Least Distance Clustering

ZSEP	Zonal Stable Election Protocol
IDSs	Intrusion Detection Systems
MTPCRP	Minimal Transmission Power Consumption Routing Protocol
DoS	Denial-of-Service
R2L	Remote to local
U2R	Unauthorized access to the root
IG	Information Gain
PCC	Pearson's correlation coefficient
PCA	Principal Component Analysis

Chapter 1

Introduction

1.1 Wireless Sensor Network

Wireless Sensor Networks (WSNs) are a group of spatially distributed autonomous sensors that are wirelessly connected to each other to monitor and collect data about the physical environment [1]. These sensors can be deployed in various areas to track the activity and then transfer the information for further processing to a special node called a sink or Base Station (BS). In WSN, a vast amount of Sensor Node (SN) analyses their surroundings, process data as well as send/receive processed information from/to other neighboring nodes such as temperature, humidity, pressure, sound, vibration, etc. Several of these SNs are dispersed around Cluster Head (CH) and are linked through the networking system to a BS in the centralized tree, mesh or other topology. The sink gathers data from the sensors (CHs or SNs) for the destination in centralized networks. In many instances, the BS is also able to turn on sensor endpoints by transmitting access control as well as monitor data through connected channels [2].

The BS in a WSN System is interconnected through the internet and utilized for data dispensation, scrutiny, withdrawal, and loading. In a clustered network design, SNs self-organize into clusters and its foundation is cluster-based leach protocol. WSNs have a group of SNs through which these nodes communicate with each other to identify useful information and data. The primary target of WSN is to detect the current event's behavior and transmit its information to the sink. Limited power and a particular battery backup are effective at controlling SNs. Due to the large distance between the SNs and the sink, more power consumption occurs in the sensors [3]. A significant obstacle to enabling a wide network transmission range is the SNs' finite energy. In the interest of improving the effectiveness and longevity of SNs for long-term operation in WSNs, the standby time and position of CHs are crucial factors. The lifetime of a WSN depends on the CH lifetime, which makes optimum use of the CH battery an important parameter for an energy-efficient WSN [4]. In designing a sensor network, the focus must be on routing path (distance between SNs to CH and CHs to BS), energy consumption, and battery lifetime as shown in Fig. 1.1. Previous research on WSNs gave different ideas to extend network lifetime by utilizing the present resources.

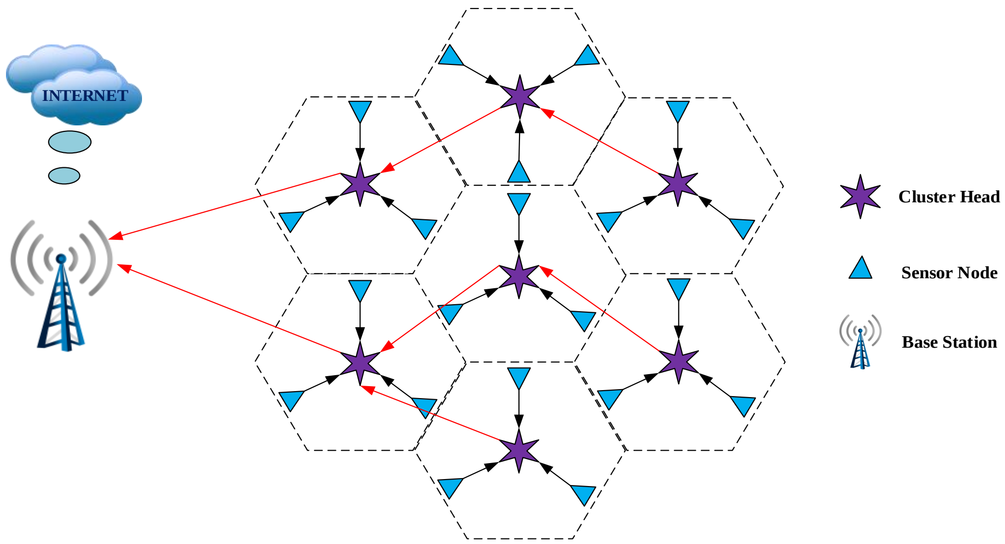


Fig. 1.1. Topology for a generic Wireless Sensor Network

Recently the advancement has opened up a whole new fascinating field for the implementation of new technologies in the energy efficiency and security of WSNs. By the next decade, the WSN is expected to be linking billions of devices, where devices interact with each other without human interference. Contact between devices is unique to the application and the network provides the internet with a backbone for application results. WSNs are an excellent option for a broad variety of applications, since they can sense the world, contact neighboring nodes, and perform simple calculations on the collected data in various circumstances. Initially, WSNs were used for military surveillance but later the applications were extended to other fields such as industrial manufacturing, health monitoring, ecosystem monitoring, and environmental monitoring. Another few applications enabled by WSNs include agricultural and vegetation monitoring, remote sensing, weather prediction, volcanic activity surveillance, military surveillance, traffic state updates, and biological monitoring systems [5]. The network market for wireless sensors was projected at US$ 46.76 billion in 2019 and is expected to hit US$ 123.93 billion by the year 2025.

1.1.1 Characteristics and Design Requirements of WSNs

The sensor network has application dependency and characteristics of WSN are highly influenced by this. The most important characteristics of WSN are summarized as follows [6]:

a) Small and low-cost devices.
b) Sensors are large, remote, densely deployed and not easy to access.
c) Sensors are with limited memory, restricted energy resources, bounded communication and limited computation storage.
d) High network dynamics and the topologies used changes very frequently.
e) Nodes are prone to breakdown.
f) High application dependency.
g) Sensor mainly used broadcast communications.
h) The data flow is between all groups or clusters and clusters to the BT and vice versa.
i) Used query dependent, periodic and event-based communication patterns.

Design of WSN is a challenging task as it is application specific and dealing low power constraints for multiple functioning of nodes. This network operates in such surroundings where no human involvement for technical support, limited sensing and computational capability.

The design requirements for WSN are as given below [7]:

a) Protocols must be scalable to handle huge number of nodes.
b) SNs should be inexpensive and reconfigurable.
c) Nodes should collect the data and work collectively for the aggregation in a systematic manner.
d) SNs must process data and also respond to change in duty assigned. So SNs should be programmable and reconfigurable during programming.

e) Guarantee for secure data communication over entire WSN specifically for deference applications.
f) The maintenance and deployment of network must be independent as no human interaction and random deployment.
g) Routing configuration algorithms should be disseminated due to fewer infrastructures.
h) Network must be adaptive to environmental changes.
i) Nodes must rely on a restricted energy production.
j) Many application scenarios can't support replacement and recharge of batteries. Noticeably, energy efficient mode of task is mandatory for the network.
k) Node failure tolerance is must as SNs can run out of energy.
l) Efficient data routing protocols must be considered for enhancement of lifetime of the network.
m) Hardware of the SNs must consider the energy constraint as its primary requirement by using the energy efficient operating platform, application software and embedded electronics chip.

1.1.2 Energy-efficient Clustering and Routing in WSN

Researchers are facing significant challenges to develop robust energy-efficient clustering and routing protocols for WSNs in different areas such as military, agriculture, education, industry, environmental monitoring, etc. WSNs have made an everlasting imprint on everyone's lives. The bulk of existing routing protocols has focused on CH election while disregarding other important aspects of routing including cluster formation, data aggregation, and security, among others. Although cluster-based routing has made a significant contribution to tackling this issue, the CH selection procedure may still be improved by integrating critical characteristics [8]. There are many performance matrices associated with the efficient working of a WSN. For an energy-efficient network, optimum use of the CH battery is essential, load balancing is important for the stability of the WSN, and the hole-free deployment of SNs is essential for the uninterrupted coverage of the surveillance region [9]. The challenge is to minimize the consumption of the energy with maximizing performance and security.

Sensor networks with constrained processing, storage, bandwidth, and energy needs additional security measures and attack prevention. Attacks may be divided into two categories: passive attacks and aggressive attacks. The passive assault is restricted to listening to and studying transmitted data. It is simpler to plan and more challenging to stop this kind of attack. Therefore, the attacker does not alter the information that is transmitted. By examining routing information, the attacker can plan an active attack by having access to sensitive data or by being aware of key network nodes (such as the CH node). An attacker attempts to delete or alter the communications sent over the network during active assaults. In order to interfere with network functionality or to inflict a denial of service, he can also inject his own traffic or repeat previous communications. We may list Tampering, Black hole, Selective forwarding, Sybil attack, Jamming, Blackmail assault, and Wormhole attack as some of the most well-known active attacks [10]. So, to enhance the energy efficiency and security issues in WSNs, we are focusing on nature-inspired algorithms.

1.1.3 Applications of WSNs

These networks are used to analyze a wide range of physical disturbances, including heat, moisture, tremors, geological occurrences, etc. They are best summed up as follows [5]:

a) **Area Monitoring:** Area monitoring is a popular WSN application in which SNs are placed throughout a territory to keep track on a certain event. For example, the sensors are used for detecting enemy intrusion as a military application and the geo-fencing of gas or oil channels as civilian applications.

b) **Health Care Monitoring:** Applications for health care monitoring range from body position measurement and location tracking to general patient monitoring in both hospitals and at home. Body-area networks may gather data about a person's fitness, wellbeing, and energy usage. Wearable and implantable medical applications might both be taken into consideration. Wearable gadgets are either worn on a human's body surface or in close contact to the user. The human body is where the implanted medical devices are placed. Both the cases need to monitor through the WSNs.

c) **Environmental Monitoring:** There are many applications in monitoring environmental constraints and provides the extra challenges to sense harsh environments.

   i. **Air Pollution Monitoring:** WSNs have been set up in many residential areas to keep an eye on the concentration of hazardous gases for occupants. Instead of using wired connections, they may benefit from wireless communications, which will make them more accessible for testing and monitoring various readings from the surrounding locations.

   ii. **Forest Fire Detection:** To find the start of a fire stroke, a network of SNs can be set up in a forest. Sensors that detect gases, temperature, and humidity that are created by fire in trees or plants can be installed in the nodes. For the firemen to successfully carry out their mission, early detection is essential and crucial. With the advantage of WSN this is possible for fire brigade to know when a fire is started and how it is spreading.

   iii. **Landslide Detection:** A landslide detection system makes use of a WSN to detect the minor changes in various parameters that may occur before or during a landslide and also to know about the activities of soil. According on the data acquired, it could be feasible to predict landslides long before they actually occur.

   iv. **Water Quality Monitoring:** Water quality monitoring involves analyzing water properties in dams, rivers, lakes and oceans, as well as underground water reserves. The installation of monitoring stations permanently in difficult-to-access sites is made possible by the use of several wireless dispersed sensors, which also allows for the production of a more accurate map of the water status without the need for human data recovery.

v. **Natural Disaster Prevention:** WSN has the power to take meaningful action to stop the effects of natural catastrophes like floods. Wireless nodes have successfully been deployed in rivers where changes of the water levels have to be monitored in real time.

d) **Industrial Monitoring:** There are many applications in monitoring industrial processes for smooth functioning of industry routine.

i. **Machine Health Monitoring:** Due to the huge cost reductions and novel functionality that WSNs provide, they have been developed for condition-based maintenance of machinery. In places where a wired system would be challenging or hard to install, such as on moving machinery and autonomous vehicles, wireless sensors can be installed.

ii. **Data Center Monitoring:** The increased concentration of compute nodes in a data center frequently causes problems with cabling and IP addresses. More and more racks are being equipped with wireless sensing devices to keep track of the intake and exhaust temperatures in order to solve this issue.

iii. **Water/Waste Water Monitoring:** Evaluating the quality of surface or subsurface water and safeguarding a nation's water infrastructure to improve the lives of people and animals are just a few of the numerous actions involved in gauging the purity and level of water. It might be applied to prevent water waste.

e) **Internet of Things (IoT):** IoT is the upcoming era of communication. According to recent reports, there will be around 9.5 billion connected devices in 2020. IoT applications are rapidly increasing all over the world but are majorly used in North America, Western Europe, and China. To expand the digital economy, IoT is the major upcoming market, according to data declares in reports. Recent and existing applications of IoT are highly promising and more efficient for the users, but high-level privacy and authentication for these applications is required. Existing solutions using different technologies like machine learning, fog computing, edge computing, and Blockchain are implemented for high-level security and privacy in IoT. The introduction of IoT-based applications demands a new algorithm to ensure the efficient working of WSNs. In WSN, different objects at different places are empowered to transmit and receive the signals generated by them. High security and privacy are required in WSN through the environment [11]. Information storage, authentication, and management are the main issues in security concerns. Different types of layers can be designed for the safety of WSN, in which the first layer includes many sensors and actuators for taking the information from different functionalities. The main security concerns that the network layer has faced are access attacks, data transit attacks, routing attacks, and unlawful attacks [12].

1.1.4 Challenges in WSNs

There are a number of security challenges, privacy concerns, reduced latency, less routing overhead, data collection and channel impairments solutions, as well as other issues and metric factors that affect the overall performance of WSN. The sensor network must take into account the following design considerations.

a) **Energy Consumption:** Due to the restricted energy supply of the sensing nodes, energy consumption is a significant problem for WSNs. To save energy, the nodes should turn off the radio power supply when not in use. As a finite energy resource controls the network's lifetime, this is required to reduce the energy use of sensing nodes. The design of the SNs depends on the type of battery used, which may cause overcharging or discharge issues due to the sensor network's low power consumption [13].

b) **Network Lifetime:** Transmission power is inversely proportional to square of the distance, network longevity and power usage are connected. In comparison to direct communication, the multi hop routing technique will therefore use less energy [14]. A sensor network's lifespan can be increased by combining many distinct strategies. For instance, energy-efficient protocols are designed to reduce the amount of energy used during network operations. But even while they are idle, node components like the CPU and radio use a lot of energy. The fixed battery capacity of the SNs, which directly affects network lifespan, should be used as effectively as possible to reduce energy consumption and, in turn, lengthen network lifetime.

c) **Data Aggregation:** With the advent of biomedical sensors, healthcare application for monitoring of vital body signs of patients is developing rapidly wherein all sensors collectively send data to the central server. It involves SNs grouping into clusters and selecting CHs. CHs aggregate data to its group and forward accumulated data to BS resulting in a higher energy spend [15]. CHs transmit data over long distances and for maintaining the smooth functioning of the network additional saving of energy and prolonging of network life is mandatory.

d) **Scalability:** Scalability is one of the foremost trials in WSN so a large number of SNs must be scalable and sufficient enough to respond to tasks in order to hundreds or thousands deployed nodes. So, the clustering and routing techniques must require different and more scalable solutions [16]. Along with scalability, adaptability and decentralization is another important standard in designing a MAC protocol. The sensor network should adapt to the changes in the topology network size, and node density. In addition, some nodes may die overtime, some may join and some nodes may move to different locations. A good MAC protocol should incorporate these deviations to the network.

e) **Cost Effective:** Considering that there are several SNs in the sensor network. Production costs must be taken into account since the cost of a single node is crucial in defending the total cost of networks.

f) **Coverage and Connectivity:** This is mandatory to handle handling vacuums and sustain certain percentage of area exposure by sensing nodes inside the WSNs along with efficient utilization of energy. In WSNs, each SN acquires a certain view of the environment and is limited both in range and in accuracy that can only cover a limited physical area of the environment. Hence, area coverage is also an important design parameter in WSNs [17]. A sensor network should also be capable of adapting to changing connectivity due to the failure of nodes or new nodes powering up. The routing protocols should also be able to dynamically consist of SNs in their paths.

g) **Security:** To achieve security is a great challenge for energy saving WSNs as it is necessary to use an algorithm that has no effect on energy resources as there is a tradeoff between security and network lifetime (sensor's energy). As SN has a limited energy, limited processing power and limited storage. So, it is essential to use security algorithm that protects the data with less energy consumption. Multiple security level must be applied on data aggregation system [18]. WSN has turned into a strong arising network innovation which has shown a precarious ascent as of late due to dependable and itemized data from any place whenever. In certain uses of WSN, it is unpredictably sent in unattended antagonistic territories like a boundary or distant districts where other than energy effectiveness, security is one more significant issue to be addressed because of the remote exchange of data. The data accumulated from SN is significant and classified for example wartime encompassing subtleties, private data of patients in clinical application, checking of full alert designs, data of traffic light framework and so on. Because of the qualities of WSN, it is exceptionally difficult to consolidate security.

h) **Quality of service:** QoS intended for WSN should be intelligent enough to support scalability. Adding or removing of the nodes should not affect the QoS of the WSN. Most of the time routing in sensor networks need to sacrifice energy efficiency to meet delivery requirements. Even though multi-hops reduce the amount of energy consumed for data collection which may slow down the packet delivery. Also, redundant data makes routing a complicated task for data aggregation affecting thus affecting QoS in WSN

[19]. Real-time communication over WSNs must be supported via provision of guarantees on minimum delay, maximum bandwidth efficient data delivery, less energy consumption, longer network lifetime etc. as QoS parameters.

i) **Latency:** The most crucial elements that must be taken into account while constructing WSN are inefficiency and data delay. Data delay is caused by data transformation and multi-hop relay [20]. Moreover, some routing protocols design their algorithms with extra overhead, which is inappropriate since it negatively impacts the energy of SNs.

j) **Synchronization:** Synchronization is a vital service in WSNs. A clock synchronization service has to provide certain problems in comparison with infrastructure-based networks [21]. The synchronization protocols initiated in existing work are Reference Broadcast Schemes (RBS) and Delay Time Synchronization (DTS) protocols. The various issues in synchronization are energy utilization of synchronization schemes and lifetime and degree of failures of SNs.

1.2 Nature-inspired Optimization Algorithms

A family of meta-heuristic algorithms referred to as "nature inspired optimization algorithms" imitates or is affected by certain natural events that are described by natural sciences. An attribute shared by all meta-heuristic algorithms inspired by nature that blend rules and randomness to mimic some natural events [22], [23]. Despite this, the SN batteries cannot be changed when installed in a remote or unsupervised area due to their wireless nature [24]. As a result, numerous types of research are being done to lengthen the life of a node span. The bulk of existing node clustering techniques suffers from non-uniform CH distribution, an imbalanced load difficulty within clusters, concerning left-out nodes, coverage area, and placement according to a recent study. Metaheuristic algorithms (GA, PSO, ACO, GWO, and BOA) have the advantages of being simple, versatile, and derivation-free, as well as effectively utilizing the network's energy resource by grouping nodes into clusters to increase the lifespan of the entire network [25–29].

This section also discusses how critical issues can be addressed by speeding up the implementation process, how more efficient data can be transferred as well as how energy consumption can be reduced by using bio-inspired hybrid optimization algorithms.

Optimization is important for solving many engineering problems in the present scenarios. Generally, optimization can be used to minimize a function subject to various equal and unequal constraints. Typically, these functions are nonlinear and very challenging to solve. There is diverse spectrum of many such algorithms including linear, quadratic, convex, conjugate, inner point, trust region and many more. Such algorithms can be used easily solving big iterative problems from traditional simple methods to evolutionary techniques and nature inspired meta-heuristic approaches. Recently, nature inspired algorithms become popular for dealing high and tough nonlinear complex issues because of their eminent properties like flexibility, higher efficiency and simplicity. These techniques primarily rely on computational intelligence, which seeks to mimic nature's plurality and distinctiveness. Though optimization is a vast area for study but there is a significant gap in-between theoretical and observational.

Mostly real-world problems can be successfully solved using practical approach of nature inspired algorithm. Though the traditional algorithms work well in performing various tasks and able to compute a wide variety of problems but very difficult to understand their functioning exactly. This is essential to have the knowledge of multi-disciplinary numerical, mathematical, dynamic system study and computational complexity analysis along with the respective tool knowledge. So, this is very challenging task. On the other hand, nature inspired algorithms have focus on investigating the essential characteristics of the technique, their elements, searching behavior and method to achieve the best solution [30]. Usually, these algorithms are characterized by initial steps, search characteristics and dynamics of algorithm. All the algorithms utilize population size or multiple search agents such as ants, bees, bats, swarms etc. Each search agent communicates to a key vector and usually has best optimum solution known as fitness objective. Multiple solutions in a population size presented diversity as well as dissimilar fitness objectives.

The development of the inhabitants is often obtained by mutation and crossover and is iterative having multiple solutions along with various properties. The measure of the exploration representatives is crisscross in the space and these moves are known as quasi deterministic movement. So, random approach is usually employed for generating new possible solutions. This random mechanics allow the solution to escape from getting stuck locally. If the search is local then there is possibility of moving locally and stuck there. On the other hand, if the search is global maximally the convergences slow down somewhat. Various algorithms have significantly different random size and behave differently for local and global search. Selection of best solutions are simply carried out by "fit to survive" so as to keep the best next generation population that is actually act as motivation for driving the miscellaneous population. Table 1.1 shows characteristics and role of nature inspired algorithms.

Table 1.1. Characteristics and Role of Nature-inspired Algorithms

Characteristics	Role
Population Size	Sample and Diverse
Randomization	Escaping local minima
Selection criteria	Driving convergence
Equations	Iterative progress of outcomes

While developing approaches and strategies for energy saving to extend the Network lifespan of WSN, it is important to consider the many concerns and obstacles that WSN faces. In wireless sensor networking, network optimization is a crucial component, and optimization methodologies are used to find the best solution to such a problem. Choosing the right algorithm is crucial in any optimization strategy because there are several optimization techniques available to meet various challenges.

1.3 Hybrid Nature-inspired Optimization Algorithms

By combining two or more algorithms according to their merits or desired qualities, hybrid optimization is a way for getting the best result possible under the given conditions. Depending on the situation, the word "optimum" can signify either the utmost or the least. The difficulties that must be taken into consideration while constructing a WSN include energy efficiency, cost, and application requirements. To make WSN effective, both the hardware and software need to be optimized. To address the various issues, numerous optimization method combinations are available.

In hybrid optimization approach, picking the right set of algorithms is crucial. Nature-inspired hybrid algorithms are designed to overcome different constraints in WSNs. Many researchers have implemented different meta-heuristic algorithms in the past to improve the lifetime, stability, and performance of the entire WSN. Hybridization techniques in optimization algorithms have helped in improving the network lifetime, stability period, throughput, number of dead nodes per iteration, and residual energy of the network. Sometimes, these bio-inspired algorithms evaluate incorrect solutions for some real-time applications. Convergence speed, multiple objective problems, dynamic problems, and local optima convergence are hot research problems nowadays. Hybridization of algorithms requires a large number of functions to be evaluated, resulting in more accuracy and improved performance of WSNs. Researchers have suggested the use of creating and optimizing a multi-objective function with a suitable mathematical function-based optimizer or hybridization technique to solve challenging, dynamic and multi-objective problems in WSN [31].

Here, we explore recently used hybridization techniques (DE-GA, GA-PSO, PSO-ACO, PSO-ABC, PSO-GWO, etc.) for bio-inspired algorithms to improve the energy efficiency of WSNs [32]–[36]. DE-GA [32] is more accurate and requires less time to complete. The technique works well in terms of accuracy and time complexity due to the rise in the population vector size. By carefully selecting the design parameters and employing superior hybrid methods, efficiency and forecast accuracy might be improved. Under the Swarm-based techniques, four different and unique techniques are listed. In terms of network lifespan and packet delivery ratio, the GA-PSO [33] method is found superior. PSO-ACO [34] hybrid optimization technique evaluates the shortest path for data transfer from the CH to the BS. The PSO-GWO [35] approach has enhanced the exploration ability by preventing PSO from falling into local minima. This hybrid technique also improved the performance of the network as compared to ABC [37], PSO, and GWO meta-heuristic methods.

1.4 Feature Selection and Classification Techniques

To enhance the performance and security of WSNs for IoT applications a novel combined feature selection method known as the Fast Correlation-Based Feature Selection (FCBFS) with Extreme Gradient Boosting (XG-Boost) has been proposed. Our framework basically comprises of dataset description, feature engineering, classification techniques and performance matrices. The most common NSL-KDD dataset is utilized for classification in the implementation. There are 41 distinct attributes in this dataset, including host type, content type, basic type, and traffic type for the selection of best features. In feature engineering, there are three categories: 1) Feature extraction 2) Feature selection 3) Feature scaling. Feature Selection (FS) helps to improve simplification by decreasing over-fitting, computational efficiency, and stuffing, as well as improving cataloging precision by developing an accurate prediction model [38]–[40]. The FS procedure is divided into four steps as discussed below:

- **Feature subset generation:** It is an exploration technique that uses one of the succeeding search strategies to pick a feature subset: full search, heuristic search, and random search.

- **Feature evaluation:** Filters, wrappers, and embedded designs are utilized to analyze the created subsets in this item.

- **Stopping criterion:** When the optimization process should be stopped, this feature determines. The determined number of repetitions or defining a goal objective, such as absolute classification accuracy, are examples of stopping criteria.

- ❑ **Validation:** When employing the created subsets in an optimization method, this procedure verifies the correctness of the results.

The classification phase is the last step in which a labeled training dataset is used for experimentation and training of the classifier with known labels. Later the effectiveness and robustness of the classifier are checked upon the testing dataset, which doesn't contain obvious labels. The training and testing stages of the model that employ classification-based algorithms can be separated into two categories. Labeled data are used in the training step to acquire a specific classifier. This classifier may then be used in the test step to determine if a test instance is normal or anomalous. Some classifiers like Decision Tree (DT), Random Forest (RF), Naïve Bays (NB), Extra Tree (ET), and XG-Boost are used in our experiment with the NSL-KDD dataset. In final step, comparison of the classifiers, performance measures such as accuracy, precision, recall, and f-score are evaluated.

1.5 Energy-efficiency and Security Requirements in WSNs

The general necessities for WSNs are given as:

1.5.1 Energy Saving Requirements for WSNs

- ❑ **Lifetime -** The network lifetime should be long lasting while consuming the energy from SNs, otherwise, there is no use of nodes after deployment.
- ❑ **Network size -** Larger network size is required in most of the applications to monitor events in a large area.
- ❑ **Minimize faults -** Short information is generated by faulty network resources. At the sensor level, the faults may cause communication link failure and loss of events in the monitoring system and sink level. In both cases, the data transmission is incomplete and therefore the collected information is not reliable in WSNs.

- **Reliability** - The network should be intelligent to deliver consistent data transmission by avoiding packet loss criteria to establish a maximum degree of efficiency in the monitoring system. Therefore, employing energy-efficient communication techniques is highly desired.

1.5.2 Security Requirements for WSNs

Transmission of sensitive information over sensor networks for claims such as military and health care put security as a major consideration for researchers. Security becomes a major risk factor because of the deployment of nodes in a hostile environment. Cryptographic algorithms deal with the secure data collection and exchange of information among the networks. Major considerations of security are memory limitations, energy constraints, hostile deployment, topology and ad hoc nature of the networks [41]. A network must involve four basic security mechanisms such as authentication, confidentiality, availability and integrity.

Attacks are the major cause of designing security protocols because an intruder will generally try to find flaws from the logical design. Also, since the WSN has only limited capacity, it is more efficient to design a security protocol for collecting and exchanging the information in constraint environments. A security algorithm must ensure access to message transmissions within the eyes of the enemies of honesty, authenticity, and trusted power [42]. Each eligible person should receive all the planned messages while ensuring the integrity of each message and the personality of the sender. Like other networks, WSNs also demand the following general security requirements [43]:

- **Integrity** - The network should ensure that any kind of message or data has not been modified during data transmission.
- **Confidentiality** - The intention of confidentiality is to protect the information between the network's innovation terminals or the sensors and the BS itself. Awareness information and routing information can be listened to, by the adversary, through eavesdropping. This sensational information can be used by the enemy for many illegal purposes, causing

severe damage such as destruction and intimidation. Hence, privacy in wireless communication channels should be provided to avoid infiltration.

- **Authenticity** - SNs are only authorized and are interested in providing information about network services. In common systems, authentication techniques check the uniqueness of the users during a connection where exceptional users can intrude. In the case of SNs, each SN and BS cannot authenticate that the data obtained was actually sent by a reliable sender or unauthorized intruder. In case false data is provided in the network, the performance of the network cannot be predetermined and most likely not expected.
- **Availability** - Access to services and data should be available at the time of their need. Due to certain attacks like denial-of-service, the device faces a number of risks and ends up losing access to the SN under attack. The lack of access influences the function of many time periods. Therefore, it is important to avoid or mitigate attacks that target the organization and provide secure routing protocols to fill the gap created by a particular attacked node.
- **Quality of Service (QoS)** - QoS-aware routing protocols consider the end-to-end delay prerequisites among sensor nodes while framing up the ways for movement stream. Each node is expected to have the same reporting rate, i.e., the quantity of packets transmitted per unit time. As a result, these protocols may not work proficiently for heterogeneous SNs. The security measures taken may influence the quality of service when computational coding is used. Security mechanisms should be light-weight so the overhead caused, for instance, by coding should be decreased and must not influence the network performance. The quality and performance of device networks involve the timely delivery to stop instance propagation of thrash and therefore increases the accuracy. In addition to these security requirements, forward and backward secrets are essential.

1.6 Research Motivation

In general, SNs have constrained storage, processing, and battery power capacities. Each node in a WSN has the ability to sense its environment, analyse the data locally, and transfer it to one or more targets via a cellular connection. SNs are subject to a number of limitations, including restricted energy sources, processing power, storage capacity, and communication range. The main issue among these is energy restriction and extending the network life. As energy ingesting determines the lifetime of WSN, it is a major issue. To solve these problems in WSNs, nature-inspired optimization methods can be used. These algorithms efficiently enhance the performance of WSNs but some issues like load balancing and security remains untouched. As sensor networks move towards depth, security is a cause of great concern in sensor network data transmission. As, SN may configure in harsh environment, a kind of encryption algorithm to realize the data transmission between nodes security has become an important area of study. The present thesis proposes hybrid nature-inspired techniques based on desired and important parameters for the above-mentioned problems. The key target of this thesis is to provide optimized concert which should satisfy the energy issues and security challenges in WSN.

1.7 Problem Formulation

Energy-efficiency and security-aware routing are the most important challenges in the wireless network. Since nodes are characterized by less energy, there is a need to transmit data with limited energy consumption. Thus, the energy utilization of the sensors in the large-scale network should be upgraded for better performance. In addition, energy balance routing should be established between the client and the receiver to lengthen the life expectancy of the system. To overcome this, the nature-inspired optimization algorithms can be used to ensure minimum power consumption and to boost the lifetime of the nodes in the network. But the selection of the group head at each interval of time during data transmission is the greatest challenge in the sensor network.

Hybrid nature-inspired algorithms are vital to boost the lifespan and security of the structure while giving the QoS essential by the system. QoS is incorporated in each segment of the network to maximize the network functionalities with energy proficient protocols. The foremost objectives of this research work are 1) Improving security, and 2) Reducing energy consumption. This thesis introduces novel hybrid optimization techniques and feature selection techniques using different classifiers to accomplish better efficiency in terms of both, security and energy. Following issues are considered while designing the multi-tier WSN.

- The SNs deployed in the network are not replaceable, so balancing the future energy of the nodes is an efficient way to improve the lifespan of the network.
- The sensed information is gathered into the destination node with the help of secure data transmission. Yet, extra energy is needed to ensure the security of communications sent through wireless channels in SNs.
- When clustering, the local changes need not affect the entire network. Thus, the optimal cluster structure reduces the data processed by nodes and data stored in sensor network nodes.
- As nodes are of high mobility (location changes occasionally), clustering formation should be carried out properly to stabilize the power of nodes. Since the nodes in the system are employed with homogeneous or heterogeneous type, the battery and the configuration characteristics should be of more concern throughout the data transmission.

1.8 Research Objectives

This primary goal of this thesis is to provide effective, dependable, and reliable solutions for energy efficiency and security in WSNs. The technique used, its application, research findings, and accomplishments for each of the research goals are listed below in this segment:

Objective 1:

- Review of the existing literature and comparison of different nature-inspired algorithms.

- To explore the key role of hybridization in bio-inspired algorithms and their comparison of the merit and demerit for energy efficient and secure WSN.
- To summarize the benefits and implementation status of different optimization algorithms and directions for future research in WSNs for IoT applications.

Objective 2:

- Design and development of a robust nature-inspired algorithm to improve the lifetime of SNs.
- Performance comparison with existing methods.

Objective 3:

- To Design and Development of an algorithm to enhance the performance and security of WSNs.
- Development of an efficient feature selection and classification technique.
- Performance comparison of the proposed framework with existing methods.

1.9 Research Contribution

1.9.1 Energy-efficiency and Load Balancing in WSNs

This thesis presents clustering and optimum routing protocol using classical and swarm-based approaches where data is directed to the BS through optimum routing path. In this area, there is a lot of research being done to find trustworthy load balancing methods. For energy efficient WSN, the thesis uses nature-inspired algorithms to evaluate the performance and security of WSNs. It is widely recognized in the scientific community that integrating many complementing characteristics derived from the applications improves accuracy and performance. It is well known that nature-inspired approaches offer reliable answers, but their implementation necessitates enormous specialized technology. The hybrid optimization algorithms GA-PSO and PSO-GWO increase the lifespan of a WSN and reduce the communication range of a SN from its respective CH and the number of hops necessary to transfer data from CH to BS. In particular, the algorithm models the lifetime of a WSN in terms of its fundamental CHs, and create fitness functions which can effectively stretch network's lifetime.

This thesis presents hybridization techniques in optimization algorithms have helped in improving the network lifetime, stability period, throughput, number of dead nodes per iteration, and residual energy of the network. Sometimes, these bio-inspired algorithms evaluate incorrect solutions for some real-time applications. Convergence speed, multiple objective problems, dynamic problems, and local optima convergence are hot research problems nowadays. Hybridization of algorithms requires a large number of functions to be evaluated, resulting in more accuracy and improved performance of WSNs. Researchers have suggested the use of creating and optimizing a multi-objective function with a suitable mathematical function-based optimizer or hybridization technique to solve challenging, dynamic and multi-objective problems in WSN. These hybrid algorithms provide deterministic CH selection, reduce the distance travelled by sensed data when being transmitted to the BS, and balance the network throughput.

1.9.2 Security in WSNs

The parameters for enhancing the network lifetime still need to be improved taking threshold value with existing recent works. This research work makes use of hybrid algorithms that are inspired by nature in order to maximize the performance and security of WSNs. These algorithms improve the network lifetime and security while simultaneously reducing energy consumption and processing delay in order to respond to environmental factors that are helpful for the targeted situation. Yet, new malware variants are discovered every day, and keeping up with the most recent patterns effectively to recognize emerging patterns in the use of credentials is another ongoing issue.

Most security methods involve more manual examination and do not take into account a wide range of properties. The earlier study effort did not report on the automated production of many features and their ideal combination, or integration at the feature level. So, future research might examine how complementary information can be extracted from big data sets. Studies must be undertaken for huge datasets that are updated with the newest features because the majority of the work is examined and reviewed on a small number of databases. The quantitative and qualitative analysis approaches were investigated separately. Because of the restricted environment, the majority of harmful behaviors are restrained. Experiments have been conducted in a restricted context, which might render any unforeseen functionality inoperable. Our study focused on security enhancement during packet transfer, clustering and routing in WSN.

1.10 Outlines of the Thesis

The thesis entitled, **"Evolutionary Algorithms for Improving Energy Efficiency and Security in Wireless Sensor Networks"** comprises five chapters followed by conclusions and future scope and a bibliography. The abstract at beginning of each chapter gives a brief outline of the work presented in that chapter. The thesis is organized as following:

1. **Chapter 1: Introduction**

   This chapter will cover the motivation and purpose of the outlined research topic. It will also contain the main idea for the development of the thesis.

2. **Chapter 2: Literature Review**

   In this chapter, a detailed study about the various nature-inspired optimization algorithms and performance metrics will be presented. This chapter will cover the state-of-the-art techniques developed in existing research work on energy efficiency and security of WSNs. area.

3. **Chapter 3: Hybrid Frameworks for Energy-efficient and Secure Wireless Sensor Networks**

This chapter contributes the comparative analysis of existing work with selective nature-inspired techniques in WSNs. It will also highlight the practicability of hybrid optimization techniques in WSNs for IoT applications and provides potential directions for further research in the area of hybrid optimization. Also, the details related to accomplish objective 1 will be mentioned in this section.

4. **Chapter 4: Design and Development of an Improved Energy-efficient Nature-inspired Algorithm in WSNs**

This chapter will highlight the discussion of the methodology adopted for developing a framework for energy efficient and secure wireless sensor networks by using an improved nature-inspired algorithm. It will illustrate the research gaps in the existing work that has necessitated the development of the research objectives. Also, all the details for achieving the 2^{nd} objective will be mentioned in this section.

5. **Chapter 5: Efficient Feature Selection and Classification Techniques in WSNs for IoT Applications**

This chapter will highlight the discussion of the methodology adopted to enhance the performance of WSNs for IoT applications using efficient feature selection and classification techniques. Previous work done on this subject and the research gaps that are existing will be illustrated. Also, the details for accomplishing the 3^{rd} objective will be discussed.

6. **Chapter 6: Conclusion & Future Directions**

This chapter will contain the brief summary of all the ideas, observations and contributions of the results obtained in each objective. Also, the future directions in each field are sketched in this section.

Chapter 2

Literature Review

2.1 Introduction

A wireless network made up of several tiny sensors and moderate transceivers can be a useful instrument for gathering information in a wide range of circumstances. The process of communicating or passing messages must be designed to preserve the inadequate energy possessions of the sensors. This chapter discusses the various clustering and routing protocols, nature-inspired optimization algorithms, and different feature selection and classification techniques that are designed to save energy consumption and security enhancement in WSNs.

2.2 Literature Survey based on Energy-efficient Clustering and Routing Protocols in WSNs

For improving the LEACH result, Wang et al. [44] introduced the LEACH-SWDN protocol. The sliding window in this protocol creates a random integer to create an interval. The CH was distributed and selected on the basis of higher residue energy via local radio transmission and also provided load balancing. The number of active nodes determines how best to dynamically modify the CH. As a result, compared to LEACH, ALEACH, and LEACH-DCHS protocols, the number of CHs is ideally distributed, and the energy is stable with the lower convention.

Manzoor et al. [45] created the Quadrature-LEACH protocol to increase the throughput, stability, and network longevity. The network is divided into four quadrants and then further into subsectors to produce optimal clusters that have better node distribution. It fixes the issue of the cluster's farthest nodes experiencing increased energy depletion. A good distribution of nodes in a certain field result from optimal clustering. In compared to the DEEC, SEP, and LEACH protocols, this protocol delivers upgraded network exposure and more efficient process.

For the purpose of offering the energy efficiency and enhanced stability period for heterogeneous sensor networks, Bara'a and Khalil [46] has presented improvements over LEACH protocol. Writers referred to it as Evolutionary Routing Protocol (ERP). The protocol achieves energy efficiency and a longer duration of stability than the HCR protocol thanks to the evolutionary approach function and increased fitness function. While energy efficiency is superior to SEP protocol, ERP fails to attain higher stability over time.

Cooperative-RPL (C-RPL) protocol, which is a cooperative and collected technique for establishing multiple instances among the nodes, was created by Barcelo et al. [47]. The coordinated effort among RPL Instances is facilitated by C-RPL. Every instance in C-RPL is designed to improve performance while using less energy. In RPL, instances and their constructors both functions apart from one another. To increase the flexibility of the routing operation, the nodes that belong to each instance must be defined beforehand. With C-RPL, nodes do not need to be predefined because they are built based on the objective function of each occurrence, their positions, and certain system conditions. A bigger C-RPL instance uses coalitions to enhance system performance but uses a lot of electricity. MATLAB is used to compare the performance of C-RPL versus RPL with diverse traffic. Simulation results show that the C-RPL protocol achieves a superior trade-off between performance parameters and energy economy.

Srivastava and Sudarshan [48] introduced an Optimized Zone-based Energy Efficient Routing Protocol (OZEEP). This technique uses two steps of clustering procedures. During the initial stage, a fuzzy inference system is employed. This system employs a fuzzy logic controller (FLC) to choose the CH by eliminating less-than-optimal nodes from contention. The best node is promoted by a genetic method to become the CH. Also, the system computes the ideal number of such CHs for acknowledging the full scope and is then updated. The initial step of screening is carried out by the fuzzy module, which identifies the nodes that can compete to become CH. GA is used in the second stage to generate an ideal, balanced distribution of CHs. OZEEP outperforms ZEEP in terms of energy efficiency and packet loss by a significant margin.

For power-aware routing, Chen and Weng [49] presented the Minimal Transmission Power Consumption Routing (MTPCR) protocol. This protocol determines the preferred routing path, which lowers energy usage. The MTPCR protocol also takes into account the fact that applications handling node mobility consume more energy, which results in a reduction in transmission bandwidth. Analyses of power consumption are obtained by keeping track of nearby nodes. By adopting a pathway maintenance mechanism, a good pathway may maintain its bandwidth. Whether or not to trigger the pathway maintenance operation depends on node density. In comparison to AODV, DSR, MMBCR, xMBCR, and PAMP, the simulation outcomes show lower power usage and improved system bandwidth.

For the goal of radiation mapping by autonomous cars, Hayes and Ali [50] devised the Proactive Highly Ambulatory Sensor Routing (PHASeR) protocol. This protocol uses the blind forwarding procedure to convey posts in a multipath manner. Collision-free TDMA MAC layer is generated by a stable time slot. The sampling frequency of the sensor, which has a predetermined length for each time slot, is used to calculate the cycle length. Blind forwarding is carried out via PHASeR using the hop-count gradient. All nearby nodes are made aware of the communication, and depending on the hop count gradient value, the receiving node independently decides if it is necessary to communicate any received information farther in the network. Scalability, traffic load, and mobility are quantitatively assessed with regard to the PHASeR approach. Better performance metrics are provided by this protocol for a variety of applications. As compared to AODV and OLSR, PHASeR uses less energy.

A hybrid energy-efficient distributed clustering protocol (HEED), also known as Multi Level HEED, was created by Singh et al. [51]. (HEEDML). The initial CH selection is done using residual energy. There is a tie when a node's range spans multiple CHs. To resolve this disagreement, intra cluster communication cost is used. Two additional criteria used for the fuzzy implementation's CH selection mechanism include distance and node density. Following the selection of CHs, data aggregation is carried out and is sent through many hops. The protocol addresses the five degrees of energy collection. Using this approach, handling uncertain situations and imperfect data is simple. In comparison to its prior iterations, the suggested protocol performs better in terms of overall energy consumption, throughput, packet transport, average latency, and transportation load.

An Energy-Aware Fisheye State Routing (EA-FSR) protocol was created by Kumar et al. [52]. Using QualNet5.0, the simulation is run with various settings. Instead of using distance as the primary criterion for choosing an adjacent node, it uses energy. In addition to the fundamental FSR mechanism, a path selection algorithm that considers energy is used. The nodes chosen based on the most energy-efficient criteria are used to forward packets. This method takes into account the distance between the nodes and the BS. Compared to FSR protocol, this protocol achieves superior performance while consuming less energy.

A particle swarm optimization-based clustering approach utilising a portable sink was presented by Wang et al. [53]. The EPMS uses virtual clustering and PSO together to achieve energy-efficient operation. To discover CHs, a virtual clustering approach makes use of the position of a node and the SNs' residual energy characteristics. The CH receives information using message-h packets. Information is sent to the sink node using the Message-s packet, which is the last one. The simulation findings show that the network's overall energy consumption has dropped to approximately 12%, resulting in a longer network lifetime than with LEACH or TTDD. Unfortunately, the jitter values and end-to-end latency have only slightly enhanced.

Artificial Fish Swarm Algorithm (AFSA), a hierarchical routing system created by Helmy et al. [54], is a SI-based optimization algorithm for reducing energy consumption and extending network lifetime. This method uses Preying, Swarming, and Following to choose the best CH. Based on the swarm behaviour, a better CH is chosen using a fitness function. In comparison to LEACH and PSO, simulation outcomes in MATLAB demonstrate the protocol's efficient functioning.

A QoS-based agent-assisted routing technique called QoS-PSO was put out by Liu et al. [55]. This technique chooses synthetic QoS, which is an adaptive PSO value that improves the performance of the entire network. Intelligent software agents looked at the communication on the network, the condition of each node's routing, and changes to the network architecture. These agents were employed for routing and network maintenance. In compared to AODV and EEABR, QoS-PSO protocol displays outstanding QoS measures for latency and packet loss. In terms of scalability, the protocol performs better for bigger networks, however massive route discovery increases network control packet overhead.

Energy-efficient routing and clustering are two of the most well-known optimization issues that Kuila and Jana [56] solved. The PSO technique was used by the authors to offer linear and non-linear formulations for these two significant optimization issues. With its multi-objective fitness function, the efficient particle encoding strategy is better suited for routing and clustering solutions that use the weighted sum approach. The transmission distance and packet forwarding latency are traded off during routing. A path from all of the gates to the base station is chosen via a PSO-based optimization method that has a somewhat shorter overall distance and fewer data forwards. In compared to GAR and MHRM protocols, the protocol performs better in terms of total data transferred to BS, network longevity, and inactive SNs.

A PSO-based Energy-efficient Cluster Head Selection (PSO-ECHS) technique was implemented by Rao et al. [57]. PSO-ECHS employs the fitness function and particle encoding. In order to construct weight function, which results in energy-efficient protocol execution, the choice of the CH is made by taking into account a number of factors, including BS distance, residual energy, and intra-cluster distance. Compared to LEACH, LEACH-C, and E-LEACH, the protocol achieves greater performance in terms of network lifespan and energy usage.

An energy-efficient hybrid Harmony Search Algorithm (HSA) and PSO-based algorithm known as HSA-PSO was created by Shankar et al. [58]. The parameters of distance and residual energy are used to pick CH. Better convergence and more search efficiency are achieved by this method. In HSA-PSO, the network parameters are initially initialized, and values are assigned to the velocity, location of the particle fitness function, and hybrid matrix. The maintenance of hybrid harmony memory, particle position, and particle velocity. In comparison to LEACH, HSA, and PSO, the hybrid HSA-PSO algorithm provides improved system lifespan, FND, LND, and standard deviation values.

An SI-based Fuzzy routing protocol (SIF) was created by Zahedi et al. [59]. Cluster creation and CH selection are the two fundamental operations in setup phase. After cluster framing, a Priority Factor (PF) is calculated using the Mamdani fuzzy framework to choose the CH. The total distance between cluster centers and instances is decreased using this approach, which also determines the number of assemblies desired for clustering. The use of fuzzy rules has an impact on how well the fuzzy scheme performs. By combining the ideas of firefly and simulated annealing, a hybrid swarm intelligence method is created for improving the fuzzy rule base. Compared to LEACH, LEACH-FL, LEACH-DT, and ALSPR, this hybrid swarm intelligence method executes balanced clustering to obtain the lowest total energy usage.

Predictive Energy Consumption Efficiency (PECE)-based routing approach was presented by Zhang et al. [60]. The suggested work in this approach is separated into phases for cluster creation and steady data transport. The first phase's CH selection criteria include residual energy, a node's degree, and the distance between nodes. PECE is organised by bee colony optimization in the subsequent stage of multi-hop data forwarding. By using less energy, the optimization structure of the method improves network performance and cluster quality while also achieving increased coverage. Compared to the conventional clustering algorithms LEACH and PEGASIS, it increases the network longevity.

IHSBEER protocol was proposed by Zeng and Dong [61]. The harmony search algorithm is easy to understand and has a powerful capacity to search globally. Adjusting the algorithm limits, initialising harmony memory (HM), updating HM, and repeating the process until the finish condition is fulfilled are crucial phases in the classic HS method. This solution solves the continuous optimization problem, although WSN routing is a discrete optimization problem due to the SNs' ambiguous behaviour. Thus, the routing issue of WSNs cannot be solved using the classic HS method. Moreover, a number of modifications to the conventional HS method are required to take into account the characteristics of the routing problem. At the beginning, it is suggested that the HM encoding be improved in accordance with the characteristics of WSN routing. The dynamic adaptation also promotes local search capabilities and prevents immaturity. Finally, an effective local search methodology used to augment the local search capacity, accuracy, and convergence speed. Comparing the protocol to the EEABR, ACORC, and EEHSBR protocols, the protocol obtains improved network lifespan.

An energy-conscious routing protocol known as PDORP with a directional transmission was created by Brar et al. [62]. This protocol combines elements of the DSR and PEGASIS routing protocols, and by combining Bacterial Foraging Optimization (BFO) with GA, the best routing path is found. Using a path-finding technique, the best route is chosen within the extensive collection of nodes that cover it. PDORP employs both the proactive and the reactive aspects of the routing concept. A sophisticated hybridization approach is employed to improve QoS and extend network lifespan. In PDORP, hybrid BFO and GA characteristics are used to improve network longevity and QoS measurements.

A pragmatic energy consumption model is utilised in the LWTC-BMA protocol, which Sahoo et al. [63]. designed to assess the network lifespan. The setup phase includes the CH selection process. A trust mechanism is built throughout the protocol to prevent any malicious node from becoming a CH. The system finds a CH that is reliable and economical with energy. For multi-hop data forwarding, the steady-state phase is used. In compared to the LEACH and

TCBMA protocols, this procedure achieves improved statistics in terms of overall energy usage, alive nodes, and energy.

A cluster-based, power-efficient routing system called ABC-SD was created by Ari et al. [64]. This protocol makes advantage of the artificial bee colony's (ABC) search capabilities, which are utilised to create low-power clusters. Using a centralised control mechanism, a linear programming formulation that makes use of a multi-objective fitness function. Clustering is done using the meta-heuristic ABC's quick convergence properties and effective features. While clustering, the CHs selected determine the paths used for inter-cluster communication. The ABC-SD method is tested with a variety of technologies in various network sizes, and the outcomes show the protocol's effectiveness in terms of coverage, packet delivery ratio, and network lifespan. Compared to PSO-C, LEACH-C, LEACH, and ABC-C procedures, it is superior.

For routing that is energy-efficient, Gajjar et al. [65] presented FAMACROW protocol that offers steady-state operation, neighbour search, and network configuration features. For CH selection, the protocol uses fuzzy logic with factors for remaining energy, network quality, and nearby nodes. Repeating the neighbour discovery stage after the steady state phase makes the FAMACROW protocol resistant to link or node failure. The FAMACROW uses ant colony optimization to quickly and accurately identify the best route from CH to MS. Uneven clustering and artificial intelligence are used to solve the hot spot problem impacting the permanent sink node. The protocol outperforms IFUC, EAUCF, UCR, and ULCA protocols in terms of efficiency, throughput, network relaxing time, and expectancy.

The Bee-Swarm protocol was created by Mann and Singh [66] and consists of three stages: building the Bee-Cluster, determining the best path, and using the Bee-Carrier to transmit data. Bee-Cluster, Bee-Search, and Bee-Carrier increase the protocol's resilience. By running the CHs selection and creation operation, the cluster is constructed. The ABC meta-heuristic is initially used in the setup phase to choose the CH. Cluster creation is completed when CH is chosen. To create the clusters, a join request is sent to each of the CH's surrounding nodes. The Bee-Search phase chooses the scout bees' routes for communication. With both forward and backward search techniques, this is maintained. The forward search is used to explore the network. The backward search strategy creates and maintains the route between various nodes and BS. Comparing Bee-Swarm to MRP and ERP, the lifetime of the network is improved by 10% to 15%.

A bee-inspired approach dubbed Bee-Sensor was introduced by Saleem et al. [67]. This protocol employs the Bee agent model and agent-to-agent communication to determine the best routes and improve protocol performance. The protocol has improved PDR, latency minimal energy use, and network lifespan outcomes. The primary traits of the various traditional and intelligence-based hierarchical routing algorithms are displayed in Table 2.1.

Table 2.1. Relevant literature based on clustering and routing protocols

Protocols	Objective	Methods	Advantages	Future Scope	Matrices
LEACH-SWDN [44]	To regulate the best CH count for energy usage dynamically.	Dynamic optimization of CHs and optimal sliding windows	Enhanced network longevity and lowered energy usage every round.	To reduce the extra network traffic brought by packet overhead.	Network lifetime, energy consumption, Optimal CH count, success rate
Q-LEACH [45]	To attain an improved stability period, network lifetime and coverage	Ideal clustering by spreading randomized clustering	Improved network exposure and energy-efficient process	Protocol expansion for large-scale networks and to shift mobility	Solidity period, network generation, and throughput
ERP [46]	To accomplish a lengthier network lifespan and stability period for networks	Evolutionary style with an adapted fitness function	Enhanced clustering and well energy efficiency, better steadiness period	Desires more perfection in fitness variant to lengthen solidity period	Average remaining energy of nodes, FND, LND, alive nodes
C-RPL[47]	To achieve healthier network energy ingesting and a fair network concert	A supportive approach and cooperative constraint α	Trade-off between energy ingesting and performance is achieved, and avoid congestion	C-RPL implementatio n for distributed network scenarios	Average number of hops, average packet delivery ratio, average energy ingesting
OZEEP [48]	To realize optimal clustering, energy effectiveness, fault lenience, scalability	Genetic fuzzy system for ideal CHs selection	It compacts energy feeding and is good for time-driven submissions	The problem of re-clustering of separated nodes to join a CH in the next round	Network lifetime and the packet drop rate

MTPCR [49]	To diminish power feasting, avoid network path damages and to recover bandwidth	Ideal path detection algorithm, path preservation mechanism	Improved network bandwidth and boosted network lifetime than MMBCR, AODV, DSR, PAMP	Augmentations for QoS consciousness	Regulate packet overhead, power feasting, hop count, average hop length, throughput, path smashing count
PHASeR [50]	To permit the mobility and strength factor in routing	Gradient maintenance measure, Blind forwarding	Low-slung overhead and improved network concert	To analyze the things of channel fading	Energy ingesting, packet delivery ratio, throughput, regulate overhead,
HEEDML [51]	To augment the network period and performance for diverse networks	Fuzzy and non-fuzzy based executions	Surge in network lifetime, PDR and low resistor overhead	Developments to handle active networks	Packet delivery ratio, throughput, average delay, and network traffic overhead
EA-FSR [52]	To guarantee the drop in overall energy consumption	Enhanced fisheye state routing, energy-aware route selection	Ensure lengthier lifetime and reduced energy consumption	Augmentation needed for high mobility-based claims	Average energy feasting, end-to-end delay, and throughput
EPMS [53]	To attain energy efficiency and improved network performance	Virtual clustering, PSO	Energy-efficient procedure and enhancement in average delay	QoS awareness	Energy consumption average delay and network lifespan
AFSA [54]	To subordinate network energy consumption	Enhanced CHs assortment using artificial fish swarm algorithm	Accomplishes better network lifetime. Less energy ingesting per round	Additional authentication of results in NS2	FND, energy consumption per round, network life time, data established by BS
QoS-PSO [55]	To increase the QoS level	PSO constructed multi agent model	Enhancement in the QoS degree	To successfully grip high mobility	Packet loss, average remaining energy, QoS, mean delay
PSO [56]	To encompass lifespan of wireless sensor networks	PSO encrypting, multi-objective fitness function, load balanced clustering	Energy feasting is balanced and network lifetime is upgraded	Improvement to deal with active network circumstances	Network life time, energy feasting, FGD, LGD, delivery of total statistics packets
PSO-ECHS [57]	To accomplish boosted network lifetime by preserving the energy of SNs	PSO, fitness function, weight function based balanced clustering	Improved outcomes in terms of total energy ingesting, and network lifetime	Fault tolerance and heterogeneous WSNs	Total energy consumption, network lifespan
HSA-PSO [58]	To conquer balanced energy feasting among SNs	Hybrid HSA and PSO	Attains better search, Fast convergence, energy-efficient	Scalability augmentation for very large-scale networks	Residual energy, mean throughput, standard deviation, FND, LND
SIF [59]	To accomplish balanced clustering and least global energy consumption	Fuzzy c-means, hybrid FA-SA (firefly algorithm and simulated annealing)	Energy-efficient and evade reservations during network process	To spread the work for large scale network topology and with mobile sink/SNs	Maximum and standard deviation of intra-cluster distance, FND, HND, LND, and success rate

PECE [60]	To lengthen network lifespan, balanced energy feasting	Bee colony optimization, Ideal clustering	Ideal cluster creation leads to higher network performance and load balancing	Mobility and Scalability of algorithm needs to be improved	Network lifetime and residual energy
IHSBEER [61]	To augment the network lifetime	Advanced harmony search algorithm	Boosted network lifetime and reduced energy consumption	To boost scalability a new model to tool on NS3	Network life time, average residual energy, standard deviation
PDORP [62]	To improve the network lifetime and upgraded QoS degree	Hybrid GA-BFO optimization, hybrid PEGASIS and DSR	Compact network resistor overhead, reckless response, good connectivity	Needs augmentation for dynamic situations	Throughput, delay, and bit error rate, network energy feasting
LWTC-BMA [63]	To extant a trust based secure and energy skilled clustering	Honey bee mating	Boosted network lifetime	Augmentation for supervision mobility	Total energy feasting, alive nodes, average residual energy
ABC-SD [64]	To design low-power scalable network and to improve energy efficiency	ABC, cost-based function	Improved network lifetime, coverage, packet delivery ratio	Heightening the deal with mobility, flexible packet frame project	Energy ingesting, first sensor dead, amount of packet delivered, LND
FAMACROW [65]	To upsurge energy efficiency and network lifespan	Inadequate clustering, ACO and fuzzy logic	Energy competent, decent scalability	To grip mobility and to add QoS consciousness	FND, LND, network settling time, and latency,
Bee-Swarm [66]	To enterprise and advance power aware procedure	ABC meta-heuristic, optimal clustering	Boosted PDR, and energy effectiveness	Request specific execution on real test bed	Packet delivery ratio, energy feasting, and throughput
Bee-Sensor [67]	To grow power-aware, scalable and concert efficient routing	Bee agent model, agent to agent communication	Minimum energy ingesting, fault tolerant behaviour	Development to knob mobility, scalability authentication	Latency, PDR, energy efficiency, lifespan, and control-overhead

Table 2.2 delivers a comprehensive instantaneous and systematic discussion on energy efficient classified routing protocols based on several performance metrics. The performance features considered for comparison are scalability, load balance, location awareness, fault tolerance, QoS, data aggregation, and energy efficiency.

Table 2.2. Summary of performance matrices for different routing protocols

Protocols	Scalability	Load Balance	Location Awareness	Fault Tolerance	QoS	Data Aggregation	Energy Efficiency
LEACH-SWDN [44]	Limited	Yes	No	No	No	Yes	Good
Q-LEACH [45]	Limited	Yes	Yes	No	No	No	Good
ERP [46]	Limited	Yes	No	No	No	Yes	Good
C-RPL [47]	Moderate	Yes	Yes	No	Yes	Yes	Good
OZEEP [48]	Very good	Yes	Yes	Yes	Yes	Yes	Very good

MTPCR [49]	Very good	Yes	No	Yes	No	No	Very good
PHASeR [50]	Very good	Yes	Yes	Yes	No	Yes	Good
HEEDML [51]	Limited	Yes	Yes	No	No	Yes	Very good
EA-FSR [52]	Good	Yes	Yes	No	No	No	Good
EPMS [53]	Limited	Yes	Yes	No	No	No	Good
AFSA [54]	Good	Yes	Yes	No	No	Yes	Good
QoS-PSO [55]	Very good	Yes	No	Yes	Yes	No	Good
PSO [56]	Good	Yes	No	No	No	Yes	Good
PSO-ECHS [57]	Good	Yes	No	No	No	Yes	Very good
HSA-PSO [58]	Moderate	Yes	No	No	No	Yes	Very good
SIF [59]	Limited	Yes	Yes	No	No	Yes	Good
PECE [60]	Limited	Yes	No	Yes	No	No	Very good
IHSBEER [61]	Moderate	Yes	No	No	No	No	Very good
PDORP [62]	Very good	Yes	No	No	Yes	Yes	Good
LWTC-BMA [63]	Good	Yes	No	No	No	Yes	Good
ABC-SD [64]	Good	Yes	No	No	Yes	Yes	Very good
FAMACROW [65]	Very good	Yes	No	No	No	Yes	Very good
Bee-Swarm [66]	Limited	Yes	No	No	No	Yes	Very good

2.3 Literature Survey based on Energy-efficient Nature-inspired Optimization Algorithms in WSNs

J. H. Holland [25] investigated the GA, a metaheuristic algorithm based on natural selection and generational reproduction of the fittest humans. Initialization, fitness, selection, cross-over, and mutation are the phases of the GA algorithm. Two extensions of GA are adaptive genetic algorithms and coarse-grained parallel genetic algorithms. They are utilized for a variety of tasks, including feature extraction, sub-set selection, engineering designs of CAD, and the traveling salesman dilemma for optimization. The two essential elements that determine the algorithm's efficiency are the fitness function and the number of iterations. By combining the parents you've chosen, you'll be able to create new genetic algorithms from the current generation (often referred to as parents) to generate offspring in the subsequent generation. They have the advantage of being able to achieve faster convergence and having a simple implementation procedure and are optimized for a wide range of functions. They have several drawbacks, such as a proclivity for concentrating on local optima rather than global optima. For decision-making issues, GA is inefficient and ineffective.

M. Dorigo [68] studied the foraging technique of ant species used in ACO. The ants leave a pheromone, a route marker that may be followed by others, to indicate a good path. A constructive greedy heuristic approach for finding excellent pathways through networks that eliminates issues. Authors use a simple phenomenon by using ant's pheromones as their path tracker by which they interact with each other. Edges and node weights are dynamically updated by the agent using the random probabilistic pheromone-based model. The used algorithm continuously repeats iterations to keep updating the path. Multiple paths are created from which the optimum path is to be selected.

D. Karaboga [69] studied an Artificial Bee Colony (ABC) algorithm, another metaheuristic algorithm that includes a food source, jobless, and employed FB. Honey bees benefit from forager feedback on food sources because it allows them to adapt and share knowledge. Employed, scout, and spectator bees are the three categories of bees. The food supply is the same as a feasible solution. The amount of nectar calculated is proportional to the solution's fitness in this algorithm. A specific operation is performed just once for a single unit of scouts, hired, and onlookers, and similarly for the other categories. D. Simon [70] studied a population-based evolutionary algorithm that, hypothetically and repeatedly, improves any mathematical function. It also improves the candidate solutions in terms of fitness function and provides a quality control technique. It is easily able to break out from local optima and obtain a fast convergence rate.

H. Shah-Hosseini [71] implemented an intelligent water drop (IWD) robust algorithm for fast convergence to the global optimization problem in WSN. In IWD, each water drop has a velocity and soil present in this journey. The velocity and the soil are determined by the quantity of soil and the time it takes to traverse the field respectively. In its course, an IWD always favours low soil content. Each IWD passes through it, producing the best solution that is utilized to update the global best solution regularly. E. Rashedi [72] presented a Gravitational Search

Algorithm (GSO) in which the active gravitational mass (GM), inertial mass (IM), location, and passive GM of each object are listed. GM and IM regulate the velocity of an object. The software navigates by modifying the mentioned masses until all of the masses are drawn to the heaviest mass, which is often considered the best choice.

X. S. Yang [73] presented a Bat Algorithm (BA) in which bats change their wavelength and rate of emission depending on how close they are to their prey. Echolocation is a method that is utilized to figure out where they are. They can tell the difference between loudness and the intensity of a pulse that should be within a specified range. A global optimization meta-heuristic approach was developed by employing bat echolocation with different pulse rates. As the bat gets closer to the location where the answer should be discovered, the frequency and strength of the pulse are modified. A. Kaveh [74] presented a Charged System Search (CSS) algorithm in which the charged particles are dispersed randomly. Well-charged particles can attract badly charged particles and vice versa. It's important to start with a modest level of investigation and gradually raise it. An examination of the global search space suggests an area in which the optimal answer is most likely to be discovered, which is subsequently exploited.

M. Clerc [75] performed an easy implementation in which the PSO algorithm can search through a huge number of potential solutions and find a suitable one over time. In a search space for candidate solutions, it also seeks the best response rather than utilizing gradients as other optimization algorithms do. The method searches the candidate solution space for the best-known solution, which is based on the particle's best-known location as well as the swarm's most advantageous position. S. Goel [76] proposed a Cuckoo Search (CS) algorithm that utilized levy flights for global search and quick convergence. Each cuckoo egg represents a fresh solution in this algorithm. The less-than-ideal cuckoos are replaced by better ones as time goes on. The eggs that survive act as solutions that are further handed down for the next iteration. With each iteration, a single algorithm iterates through the solutions in the nest as called the search space, which helps in enhancing the quality of the solutions. Its purpose is to develop better and fresh ideas, the quality of which is determined by an objective function that is often maximized.

X. S. Yang [77] presented a robust Flower Pollination Algorithm (FPA) using levy flights, cross-pollinators do global pollination, while local pollination is comparable to local search. Floral similarity influences reproduction with the fittest surviving and reproducing optimally in terms of numbers and fitness. An iterative approach that uses local and global pollination to find the best result. A. Sabry Eesa [78] studied the Cuttlefish Optimization Algorithm (COA) in which reflection and visibility are achieved by using the multiple layers of the fish. Pattern matching is simulated by visibility, whereas matching light is simulated by reflection. The algorithm uses reflection and visibility to try to hide the fish in the surroundings, and the resulting pattern is the global optimal solution. Global search with a random component is represented by the first two solution groups, while local search and solution comparison is represented by the last two.

S. Mirjalili [28] presented a Gray Wolf Optimization (GWO) algorithm that equilibriums the state of exploration and exploitation. It consists of a hierarchy of wolves i.e., alpha, beta, delta, and omega. The three finest solutions will always lead you to the ideal search space, but we must strike a balance between exploration and exploitation. The basis of an optimization technique is a series of randomly generated solutions, each of which is a vector of parameter values. Encircling the prey and pursuing the target in the search space in an iterative cycle to find the global optimum. Mirjalili [79] emphasized exploration via global search. Every quest is an attempt to find the neighbourhood's best solution. Encircling the prey is simulated by updating the location vector. The goal function determines the convergence behaviour. Exploration of the search space for the best potential answer is based on the cycled position of search agents.

S. Gao [80] presented a divided algorithm called as Improved Artificial Fish Swarm (IAFS) algorithm. It is based on searching a mathematical function, its swarming and chasing through various means, and finally its leaping behaviour.

Random behaviour is strongly influenced by the visual scope. Swarming takes place only when the current function value is better than the prior one. One algorithm iterates repeatedly and updates the swarming behaviour. To discover the global optimum or optimal solution, a fish simulation behaviour is done through a randomized parallel algorithm. Y. Y. Hao [81] proposed an improved Glow-worm swarm optimization (IGSO) algorithm to enhance the performance of multi-dimensional problems and convergence rate. A population of glow-worms with an identical quantity of luciferin is dispersed across the search space. The value of luciferin is determined by the glow-worm's location. Brighter light in the field indicates more amount of luciferin. To update the position of the glow-worm, a randomized algorithm based on parameter adaptation is used which is followed by the luciferin update.

T. Shankar [58] proposed a Harmony Search Algorithm (HSA) PSO-based hybrid algorithm to overcome the different local search constraints and exploration-exploitation trade-offs. This hybrid scenario also obtained a fast convergence rate in global search and has improved the lifetime of SNs due to its dynamic capability and high search efficiency. S. Su [82] presented a GA-PSO hybrid approach that explores the distributed clustering levels for large-scale WSNs. At the lower level, GA is used in independent subgroups for a global search and at the upper level, the PSO algorithm is used for the local search of individuals. The proposed technique also reduces energy consumption and accelerates convergence speed.

J. Kapoor [83] proposed an improved protocol of LEACH which is inspired by Low-energy adaptive clustering with the collaboration of GA and Bacteria Foraging (BF) implementation to overcome the disadvantages of former conventional protocols. The proposed algorithms also reduce energy dissipation and improve the network's lifetime. B. Farnad [84] presented a new hybrid approach by combining GA, PSO, and Symbiotic Organisms Search (SOS) based on the natural selection phenomenon. GA creates and picks the best population for the future phases, PSO accumulates and updates experience for each appropriate solution, and SOS builds on prior phases and executes symbiotic interaction update phases in the real-world population. S. Potthuri [85] proposed DE-SA hybrid approach for the best CH selection. It is utilized to extend the life of the network by delaying the death of CHs. The cluster head's remaining energy and the distance between the nodes are taken into account by the fitness function concerning the differential evolution and simulated annealing concept (DESA). In this technique, the authors have tried to keep the maximum no. of SNs alive, as the network's lifetime is directly proportional to the number of nodes alive.

Table 2.3. A comprehensive comparison of nature-inspired algorithms in WSNs

Algorithm	Advantages	Nature of Solution	Disadvantages	Applications
Genetic Algorithm (GA) (J. H. Holland 1992) [25]	Enables us to explore a search space without losing partial solutions. Explores various parts of the solution space simultaneously. Effectively combines novel combinations with existing information.	Based on natural selection and generational reproduction of the fittest humans. Initialization, fitness, selection, cross-over, and mutation are all phases in the algorithm. Follows a constructive greedy heuristic approach for finding excellent pathways.	More computational time. Large complexity in network structure. Difficulty in showing branching and looping. For decision-making issues, genetic algorithms are inefficient and ineffective.	Data clustering and mining, Traveling salesman problem, Neural networks Wireless sensor networks, Medical Science, Vehicle routing problem.
Particle swarm optimization (PSO) (Kennedy and Eberhart 1995) [75]	PSO has a few parameters to adjust, it takes less amount of time to compute, In global optima, PSO has a greater likelihood and efficiency.	The algorithm explores the candidate solution space for the best-known solution and is based on the best location of the entire swarm and the particle's most-known position.	The challenges faced in PSO are low-quality solutions, Initial design parameters are quite difficult to define, and can't to perform in the problems of a non-coordinated system.	Price and load prognostication, Volatile power management, Ideal power flow, Neuronal network training.
Ant colony optimization (ACO) (Dorigo et al. 1996) [68]	ACO can be used in dynamic applications, it can adjust to new distances and other changes, and It may also search for a large population at the same time.	Model-based on random and probabilistic pheromones, the agent dynamically updates the weight of nodes and edges (ant).	The probability distribution might change with each iteration, ACO's theoretical analysis is complex, unknown time to reach a point of convergence.	Job shop forecast problem, Retro vehicle steering problem, Image dispensation.
Artificial bee Colony (ABC) Algorithm (Karaboga 2005) [69]	ABC has a simple structure, It also uses a few parameters, Strong robustness.	The nutrition supply is comparable to a potential resolution, and the volume of nectar is proportional to the solution's suitability.	The challenges in ABC are the slow speed of convergence and low QoS, and the precision of optimization is low.	Image Processing, Clustering & mining, Fiscal communication problems.
Biogeography based optimization (BBO) (Simon 2008) [70]	Convergence occurs quickly, and can easily break out from local optimum conditions.	Species extinction, migration, and speciation are all factors that influence the evolution of species.	Exploration of the solution space is inadequate, there is no way to save each generation's finest work.	Traveling Salesman Problems, Feature Extraction.

Intelligent water drop (IWD) (Shah- Hosseini 2009) [71]	Simple to use, Union to the comprehensive optimum is a foregone conclusion, Robust	In terms of minimum direction and maximum velocity, we're looking for a global optimum.	For big enough iterations, the best solution is discovered, and Probability is used to choose the next node.	Traveling Salesman Problems, Multiple knapsacks, Workflow scheduling
Gravitational search algorithm (GSA) (Rashedi et al. 2009) [72]	Adjustable learning rate, an algorithm with less memory, Result is more consistent and precise.	The program navigates by modifying gravitational and inertial mass which is the best option.	Intensely computational, initial population and its size have an impact on performance.	Renewable micro-grid, Regulator strategy, Wireless sensor networks.
Bat algorithm (BA) (Yang 2010) [73]	Switching from exploration to exploitation is possible, offering control over parameters.	The echo sounding of bats with different beat rates was used in a global optimization metaheuristic method.	A large variety of objective functions is available, but the pace of convergence is slow. Initial parameters have no values specified,	Clustering cataloging, Facts withdrawal, Image processing.
Charged system search (CSS) (Kaveh and Talatahari 2010) [74]	During exploration, it is difficult to become imprisoned in the local minimum, Simple to implement, Initially, there are just a few settings to tweak.	An ideal explanation for optimization that mimics electrostatic services between particles and their dependency on distance.	The computation cost rises as the quantity of charged particles rises, only a few charged particles are required for preemptive convergence to occur.	Water dispersal networks, Operational mutilation recognition.
Cuckoo Search (CS) (Goel et al. 2011) [76]	Convergence occurs quickly, Global optimums are guaranteed if enough time is given.	Its purpose is to develop novel and better solutions, the quality of which is determined by an objective function	It's possible to get stuck in a local optimum along the border, Lack of effectiveness.	Exercise of neural system, Conniving a wind turbine edge, Synthesis in WSNs
Flower Pollination (FPA) (Yang 2012) [77]	Easily able to break out from local minima, the Fitness function made it easy to survive,	Optimal reproduction and survival of the fittest in terms of numbers and fitness.	Flights on Levy might lead to domain exploration and there are no default values for the initial parameters.	Design pressure pitchers, Image firmness, and Chart coloring.
Cuttlefish optimization (CFO) Algorithm (A. Sabry Eesa 2013) [78]	Can easily break out from local optimum conditions, Ensured global optimal location.	Two of the solution groups are for global search with a random component, while the other two are for local search with a solution.	Intensely computational, Slow conjunctional.	Control systems, Information mining, Biomedical engineering.
Grey wolf optimizer (GWO) (Mirajlili et al. 2014) [28]	Simple to contrivance, Flexible, Mountable, Exploration, and extraction are in a state of equilibrium.	A collection of random solutions is used in the optimization process, with each result being a vector that reflects the parameter values.	Only single-objective issues are allowed, Local optima stagnation occurs when there are a large number of variables.	Design and alteration of controllers, Clustering, and Robotics.
Whale optimization algorithm (WOA) (Mirjalili and Lewis 2016) [79]	Exploration via global search is emphasized by the global optimizer, there are fewer settings, Simple to use.	Starts with a random solution, then updates its position using a randomly selected search engine or the best approach so far.	Low rate of merging, Low precision, and Randomness affects the convergence phase.	Workflow planning of creation sites, Image segmentation, Optimal power flow problem
H-HSA PSO Algorithm (T. Shankar 2016) [58]	Searches at a higher rate, allowing for speedier exploration and exploitation, dynamic capacity.	At each level, the starting settings are modified. The goal is to provide an energy-efficient CH selection	Convergence rates start to drop in high-dimensional problems, and it's tough to fine-tune the basic settings and its full potential.	Feature selection, Training neural networks, Economic Dispatch problems.

H-GA PSO Algorithm (Shengchao Su 2017) [82]	In large-scale WSNs, it is effective for distributed clustering, and the convergence speed will be increased.	The algorithm looks for the most well-known solution in the candidate solution space and is determined by the particle's finest position.	Dependent on the initial parameters such as location, and it's possible to fall in local optima in high dimensional space.	Clustering, robotics Neural network training.
H-GA BFO Algorithm (J. Kapoor 2017) [83]	Obtained optimal coverage with less no. of nodes in large-scale WSNs, reduced avg. power consumption.	Energy and Physical parameter are to be initialized, during iterations, if the node is dead. The route is updated after one algorithm repeats repeatedly.	Dependency of fitness functions on various parameters, Difficult to implement on large scale WSNs.	Biomedical engineering, Wireless sensor networks.
Improved Artificial fish swarm (IAFS) (S. Gao 2018) [80]	Ability to make a proper junction, Suppleness, Present with great precision and fault tolerance.	A randomized parallel method that models fish behavior to get the worldwide finest or topmost solution.	Each fish's visual range is unique and can't be generalized, there is a discrepancy between global and local minima.	Job scheduling, Image processing, Clustering.
Improved Glowworm swarm optimization (IGSO) (Y. Y. Hao 2018) [81]	For many peaks, an adaptive local judgment is made, for issues with a constant domain, this method works well.	To update position, a distributed algorithm based on luciferin apprise uses the statistics accessible in the nearby vicinity.	For high-dimensional issues, performance is poor, convergence occurs gradually, and Inadequate local search capability.	Positioning numerous mobile signal bases, WSNs, Communal transport report system.
H-GA PSO SOS Algorithm (B. Farnad 2018) [84]	In the real-world population, it runs symbiotic interactions to update stages. Superior in terms of convergence and execution speed.	Searching the solution with logarithmic spirals which is a deterministic dynamical system. Natural selection inspired the notion of merging three algorithms.	Improve search performance by introducing randomization, Initial parameter selection has a significant impact on performance.	Job scheduling, Data mining, Path planning, Statistics synthesis in wireless sensor networks.
H-DE SA Algorithm (S. Potthuri 2018) [85]	Extend the lifespan of the network by extending the CHs' death, High selection rate.	Iteratively improves a potential solution based on an evolutionary process to optimize a problem.	Low convergence rate, randomness during the selection of initial parameters, and Less robustness.	Multidimensional global optimization, Training of neural networks.
H-FF PSO Algorithm (B. Pitchaimanickam 2020) [86]	Flexible, Scalable, providing parameter control statistics, Not easily trapped in the local minima.	Set the FF and PSO parameters to their default values. Calculate your fitness level and update the velocity and position of the population.	Initial characteristics such as position and velocity showed randomness. Difficult to implement on large scale WSNs.	Automatic data clustering, Machining parameter optimization,
F-GWO SFO Algorithm (L. Nagarajan 2021) [87]	Automates the setting of a collection of parameters in such a way that the weight is evenly distributed, it also enhances the stability.	An array of random solutions is used in the optimization process. The energy consumption and separation distance are considered for selecting optimal CHs.	Low solving precision. In a high-dimensional space, it's possible to fall into the trap of local optima.	Engineering Design Problems, Design and controllers tuning, Robotic and path planning.

Table 2.3 provides a comprehensive comparison of algorithms based on their strengths, drawbacks, performance-influencing variables, and application areas.

B. Pitchaimanickam [86] proposed Firefly Algorithm with Particle Swarm Optimization (HFAPSO) which is provided in the LEACH-C approach for selecting the optimal CH selection. The hybrid method optimizes firefly's global search behaviour and achieves optimal CH positioning using PSO. The suggested methodology's performance is measured by the number of live nodes, available energy, and performance. L. Nagarajan [87] proposed a Hybrid GWO-based Sunflower optimization (HGWSFO) technique for optimum CHS under particular factor constraints such as energy consumption and separation distance, to extend network longevity. Balancing the exploration-exploitation trade-off increases network performance in terms of total throughput, node residual energy, dead nodes, alive nodes, network survivability index, and convergence rate.

2.4 Literature Survey based on Security in WSNs

Authors in [88], [89] presented a Hybrid Intrusion Detection System (HIDS) that comprises a module for detecting anomalies and abuse. The objective is to leverage the profits of abuse detection and variance detection to increase recognition rates while lowering false-positive rates. However, to integrate the detection findings and describe the sorts of assaults, a decision-making module is employed. With hybrid detection, this technology not only reduces the risk of an attack in the system but also helps the user manage and fix the system. The simulation results show how well this approach works: it has a detection rate of 99.81 percent, a false positive rate of only 0.57 percent, and an accuracy of 99.75 percent.

Authors in [90], [91] identified anomalous intrusions with an intelligent algorithm, and decision criteria for FS are provided. Support Vector Machine (SVM) and Simulated Annealing (SA) may be used in the proposed technique to discover the high-ranked features to advance the accuracy of anomalous intrusion detection. Decision Tree (DT) and SA can gain decision rules for novel assaults and enhance classification accuracy by studying information from the KDD'99 dataset. This research supports the advancement of anomalous intrusion detection. The first step is to pick features for anomalous incursion detection. The second need is to use the dataset's decision criteria to detect new threats.

Authors in [92], [93] gave several weights to each feature and suggested a multi-measure multi-weight feature identification strategy that combines the filter and wrapper FS methods with clustering algorithms. When comparing the new technique to current algorithms utilizing four classifiers, it is clear that the suggested strategy outperforms the existing algorithms in the vast majority of situations. The two types of characteristics, significant and somewhat important, may be utilized to model the assault detection system to improve detection accuracy while decreasing detection time. Authors in [94], [95] suggested two intrusion detection algorithms: one employs an ensemble learning mechanism at the SN level, while the other employs an unlabelled data mechanism at the sink node and base station levels. The algorithms provide a binary tree-structured collection of detection rules as a result of their work. In comparison to prior work for intrusion detection, the new methods delivered good detection accuracy with a smaller number of chosen characteristics, reducing complexity and computation. In comparison to existing ways, the proposed methods reduced the complexity of decision-making and reduced energy use. When applied to various algorithms, the selected characteristics of the introduced algorithms produced competitive results.

Authors in [96], [97] presented a faster Deep Neural Network (DNN) architecture for detecting anomalies in network data. The NSL-KDD dataset is used to calculate the training duration and evaluate the detection mechanism's efficacy. These studies focus on using an accelerator platform to handle large datasets and detecting complicated relationships in the input dataset to identify distinct attack types. One of the major goals of these researches is to compare the act of DNN training on diverse machine types and core counts. Authors in [98], [99] proposed a customized dataset for WSN to aid in the exposure and organization of four types of DoS attacks: Blackhole, Gray-hole, Flooding, and Scheduling assaults. This research looks at the LEACH protocol, which is one of the most extensively used cluster-based routing methods in WSNs. Using technology, data from Network Simulator 2 (NS-2) was gathered and processed to give 23 characteristics. This study emphasizes the need to address security early in the network protocol development process. Without this, hostile attacks would progressively target inherited weaknesses in various network protocols and other software.

Authors in [100], [101] proposed a Restricted Boltzmann Machine (RBM) and a deep belief network to build a deep learning strategy for anomaly identification. To conduct unsupervised feature reduction, we employ a one-hidden layer RBM. This RBM's weights are transmitted to another RBM, resulting in a deep belief network. A fine-tuning layer that contains a Logistic Regression (LR) classifier with multi-class soft-max feeds the pre-trained weights. We developed a multi-class Logistic Regression layer and trained using 10 epochs on the enhanced pre-trained data to improve the network's overall performance. We created a network with a fast-training period and little dataset pre-processing. Authors in [102], [103] verified a range of ML methods, including neural networks, that leverage IoT-specific network characteristics to influence FS that may result in high accuracy DDoS detection in IoT network data. We demonstrated in this paper that ML at the packet level is possible. DoS detection from consumer IoT devices can accurately discriminate between regular and DoS attack traffic. We used a limited feature set to lower the computational cost, which is crucial for real-time classification and middlebox deployment.

Authors in [104], [105] based on the C.4.5 algorithm, the authors developed a DDoS detection system to combat the attacks. When paired with signature detection techniques, this strategy creates a decision tree that can automatically and effectively identify signature assaults for DDoS flooding attacks. A comparison of several ML algorithms and methodologies was offered. We concluded from the data that our suggested solution, which uses the C4.5 algorithm to perceive DDoS assaults, offers more precise outcomes than existing ML algorithms. Authors in [106], [107] suggested a layered hierarchical approach for identifying abnormalities in various online social networks. The experiment was conducted using four data sets, with Twitter and Facebook showing the best outcomes in anomaly detection. Existing solutions either focus on a fixed collection of named data or monitor the behaviour of randomly selected hubs rather than the unstructured flow of data in interpersonal groups. Additionally, the characteristics employed in the mining process might be expanded to provide more precise findings.

Authors in [108], [109] created a clustering-based approach for separating data describing network traffic patterns, including both normal and DDoS activities. The characteristics are used to identify assaults at the victim's end, and the process is shown with three features that may be observed in the target system. To label the data and establish classes to separate assaults from normal traffic, a voting approach is also offered. Using a demonstrative fraction of the benchmark CICIDS2017 dataset and additional vectors of attack, the technique is also verified for label assignment accuracy. Authors in [110], [111] provided two tactics in an effort to close a deficit in the field. First, name/address solutions that locally assign distinct addresses to sensor devices and then reuse them in a spatial way. Second, name/address solutions that assign location-insensitive spanning-tree topology to name/address reuse in event-driven WSNs. Additionally, the proposed methods solve network-level uniqueness and create a roadmap with smarter and better protection measures throughout the network.

Table 2.4. A comparison of feature selection techniques with different classifiers

Author	Dataset	Feature Selection Method	Number of Features	Attacks Detected	Detection Methods	Result Analysis
Norouzian et al. [112]	KDD CUP 99	Feature Vector	13	DoS, R2L, U2R, Probe	ANN	Classification Accuracy: 96.78%, FAR: 2%
Lin et al. [90]	KDD CUP 99	Decision Rules	23	DoS, R2L, U2R, Probe	SVM, DT	Accuracy with 41 features SVM: 99.03%, DT: 98.5% Accuracy with 23 features SVM: 99.96% DT: 99.6%
Mukherjee et al. [113]	NSL-KDD	CFS, IG, Gain Ratio, FVBRM	CFS:10, IG:14, Gain Ratio:20, FVBRM:24	DoS, R2L, U2R, Probe	NB	Accuracy with 10 features: 97.5% Accuracy with 14 features: 95.3% Accuracy with 20 features: 95.2% Accuracy with 24 features: 97.7%
Chae et al. [114]	NSL-KDD	CFS, IG, Gain Ratio, Attribute Ratio	Attribute Ratio:22, CFS:25, IG:23, Gain Ratio:19	DoS, R2L, U2R, Probe	DT	Accuracy with 22 features: 99.7% Accuracy with 25 features: 99.8% Accuracy with 23 features: 99.8% Accuracy with 19 features: 99.9%
Thaseen et al. [115]	NSL-KDD	PCA	10	DoS, R2L, U2R, Probe	SVM	DR: 99.4% FAR: 0.0015
Subba et al. [116]	NSL-KDD	PCA	17	DoS, R2L, U2R, Probe	SVM, MLP, C4.5, NB	Accuracy and DR for 17 features: SVM: 99.13% and 98.68% MLP: 96.76% and 96.13% C4.5: 96.85% and 97.23%

Osanaiye et al. [117]	NSL-KDD	Ensemble Multi-filter	IG:14, Gain Ratio:14, ReliefF:14, EMFFS:13	DoS, R2L, U2R, Probe	J48	Accuracy: 99.67%, DR: 99.76%, FAR: 0.42
Kumar et al. [118]	NSL-KDD	CFS, Gain ratio,	CFS:10, Gain Ratio:18,	DoS, R2L, U2R, Probe	NB	CFS with 10 features: Accuracy: 98.0% FAR: 0.101 Precision: 0.96 F-score: 0.97 Gain ratio with 18 features: Accuracy: 98.6% FAR: 0.002 Precision: 0.98 F-score: 0.99
Thaseen et al. [119]	NSL-KDD	Chi-Square	31	DoS, R2L, U2R, Probe	SVM, RBF	The average result analysis for all classes follows: TPR:0.987, FPR:0.001, Precision:0.995, Recall:0.996, F-score:0.996
Meftah et al. [120]	UNSW-NB15	RFE	5	DoS, exploits, backdoor	DT, NB, SVM	Accuracy (DT): 85.4% Accuracy (NB): 60.70% Accuracy (SVM): 70.21%
Khraisat et al. [121]	NSL-KDD, AFDA	Not Applicable	All features	DoS, R2L, U2R, Probe	SVM	Accuracy for NSL-KDD: 93.24% Accuracy for AFDA: 97.04%
Bhati et al. [122]	NSL-KDD	Not Applicable	All features	DoS, R2L, U2R, Probe	SVM	Accuracy for Linear SVM: 96.1% Acc. for Quadratic SVM: 98.6% Acc. for Gaussian SVM: 98.7%
Krishnaveni et al. [123]	NSL-KDD	IG	10	DoS, R2L, U2R, Probe	RF	Accuracy: 96.34%
Thakkar et al. [124]	NSL-KDD	Chi-Square, RFE, IG	Chi-Square: 10, RFE:13, IG:18	DoS, R2L, U2R, Probe	DT, RF, LR, k-NN, SVM, NB, ANN	SVM with RFE outperformed other classifiers.

2.5 Chapter Summery and Gaps in the Study

In this chapter, the problem of clustering is investigated thoroughly and extensive literature has been reviewed. Initially, the classical techniques have been studied and explored various existing techniques for clustering and routing in WSN. Thereafter, optimization algorithms have been studied for exploring the same issue. Traditional approaches have not been successful in locating solutions to challenging or complicated issues. Nature-inspired algorithms are heuristic approaches to problem-solving that draw on the actions of nature. The hybrid nature-inspired algorithm's architecture exemplifies good issue formulation, fitness function-based assessment of solution quality, and generation of a fresh set of solutions via new operator definitions. The major goal is to use various nature-inspired algorithms in WSNs to create energy-efficient networks with high throughput, reduced packet loss, reduced energy consumption, and longer network lifetime.

The highlights of the gaps in the literature are enlisted below:

- Researchers have used conventional clustering algorithms to achieve energy-efficient WSN by taking specific distance as a parameter.
- Fewer research work in hybridization techniques to improve the efficiency of WSNs.
- Need for load balancing to enhance the lifetime of CHs and its effect on WSNs.
- Fewer research works on security threats during attack detection and malicious node identification in WSNs.

2.6 Conclusion

Numerous approaches for energy-efficient and secure WSN have been studied in the literature. Various clustering and routing protocols considered in the literature to enhance the security and lifespan but none of them has been well thought-out effective transmission distance parameter yet. The distance from SN to BS has not been taken into account if the SN present in the network is nearer than CH. These works are restricted to specific distance parameters only, which makes the problem appearance more costly and unbalanced power distribution in WSN. Due to load unbalancing, SNs will be depleted quickly which results early breakdown of WSN. To overcome the above problems authors focused on optimization techniques. Different types of nature-inspired optimization algorithms have also been considered in the literature to enhance to efficiency of WSNs. Later, authors also considered hybrid techniques in optimization algorithms to improve the lifetime of WSNs.

Various attacks like sybil, blackhole, worm-hole, flooding, and sinkhole, security in WSNs should be considered as an important aspect. As the next era of communication is IoT, current studies have various drawbacks related to the security. Therefore, authors need to be focused towards effective security techniques to prevent the WSNs from different type of attacks.

Chapter 3

Hybrid Frameworks for Energy-efficient and Secure Wireless Sensor Networks

3.1 Introduction

WSNs have a group of SNs through which these nodes communicate with each other to identify useful information and data. The primary target of WSN is to detect the current event's behavior and transmit its information to the sink. This distribution phenomenon protects against data loss in a specific coverage area. Due to the large distance, more consumption of power takes place in the sensors. The limited energy of SNs is a major drawback to empowering a large network coverage area. The lifetime of a WSN depends on the CH lifetime, which makes optimum use of the CH battery an important parameter for an energy-efficient WSN.

In designing a sensor network, the focus must be on utilizing energy consumption and battery lifetime. Previous research on WSN gave different ideas to extend network lifetime by utilizing the present resources. Clustering and routing algorithms also have a specific contribution to the process. There are many performance matrices associated with the efficient working of a WSN. For an energy-efficient network, optimum use of the CH battery is essential, load balancing is important for the stability of the WSN, and the hole-free deployment of SNs is essential for the uninterrupted coverage of the surveillance region.

The introduction of IoT-based applications demands a new algorithm to ensure the efficient working of WSNs. In WSN, different objects at different places are empowered to transmit and receive the signals generated by them. High security and privacy are required in WSN through the environment. Information storage, authentication, and management are the main issues in security concerns. Different types of layers can be designed for the safety of WSN, in which the first layer includes many sensors and actuators for taking the information from different functionalities. The main security concerns that the network layer has faced are access attacks, data transit attacks, routing attacks, and unlawful attacks.

3.2 Energy Efficiency in WSN

The selection of CHs randomly from SNs with less residual energy is defined by LEACH[125]. There have been other differences in LEACH with respect to WSN performance enhancement. The fuzzy logic-based clustering algorithm in [126] contains three fuzzy node energy, centrality, and concentration variables. But a big problem faced by this logic is the involvement of a single cluster in the process. Authors in [127] worked on energy dissipation of the network, but the distance parameter is not mentioned. This avoidance of the distance parameter creates unbalancing in the network.

Authors in [128] considered only one parameter node centrality to minimize energy consumption. Other important parameters like load balancing, transmission distance, and uniformity in deployment are not considered. A novel algorithm in [129] is presented to minimize the distance between sink and CH. The transmission distance is specified in this algorithm, but the communication distance between SNs and CHs is not considered.

To enhance the lifetime of a WSN, a Partial Swarm Optimization (PSO) in [56] is used. In PSO, they consider the fitness function parameters like network lifetime and CH to sink distance to achieve an energy-efficient network. But the SN to CH distance is not considered in PSO. Energy description during the process is also not considered. Authors in [130] proposed a significant algorithm to tackle optimization problems. This algorithm is highly flexible and more desirable for automatically configured techniques. In complex applications, ACO can be used. Changes such as new distances may be adjusted. It can also search among a population in parallel. The challenges faced are probability distribution changes for each iteration. ACO has a difficult theoretical analysis and uncertain time to convergence.

A new nature influenced by the authors in [131] suggested a technique of Optimization Based on Biogeography (BBO-C) in WSN for energy-efficient clustering. To maximize the performance of the network, a complex and effective fitness feature is used. Authors in [132] proposed a clustering and routing algorithm (EBBO-CR) for extending the lifetime of WSN. It also minimizes the distance from its respective CHs and the number of hopes required to move data from CH to sink. But in exploiting the solutions, BBO and EBBO are bad. They have no provision for selecting each generation's best members and the resulting fitness solution.

A novel version of BBO was proposed by the authors in [133]. The implementation of the age-structured BBO (ASBBO) is the impact of people's age on the migration method. Individuals are implemented to capture age-related trends of birth rates and survival rates. Age systems enable information to be shared not only on the basis of the present goodness. In [29], the writers mimic the leadership structure and hunting method of grey wolves. Four forms of the grey wolf were used to determine their fitness, i.e., alpha, beta, delta, and omega. For the energy-efficient WSNs and load balancing of the CHs, they suggested DBSCDS-GWO and DBSC-GWO. Yeah, but the author does not take WSN's accuracy and convergence.

3.3 Load Balancing in WSN

Recent researches described many algorithms to enhance the performance of a network. In [134], the authors described a breadth-first tree using a clustering algorithm to find the gateway which is at low load. The drawback of this algorithm is the time of execution while calculating large computations. Authors in [135] build a load-balanced CDS using GA and load balanced allocation of a SN to CH. Before this, a limited study has been done to achieve a load-balanced CDS WSN.

The problem that is to be observed in this algorithm is high energy dissipation and slow convergence rate.

For congestion management of wireless sensor networks, a load balancing protocol inspired by GA [136] has been introduced. A fuzzy-based PSO routing algorithm was proposed by the authors in [137] to balance the load and scalability of the network. The authors improved three important factors like density, energy consumption, and transmission distance by using FBPSO. They also balanced the CHs' energy dissipation and increased the lifespan of WSNs. The choice of CH and network stability was not considered by them.

Authors in [138] proposed a data dissemination strategy using nature-inspired Ant Colony Optimization (ACO) named TMLBS. This paper identified three load balancing systems, which are the system of load decentralization, the system of maintenance, and the system of load diversion. The authors tried to create multiple transmission paths for the entire network at different places. In these schemes, they implemented multipath subtrees to avoid excessive load, pheromone update mechanism, and paths with low data load, respectively. But they didn't focus on the lifetime of the network and convergence speed. New nature-inspired techniques have been proposed by authors in [9] Biogeography Based Optimization (BBO-C and BBO-CR) for WSN energy production. This leads to better sensor distribution and a well-balanced method of clustering. But there is no provision for the BBO-C and BBO-CR to pick the best members from each generation. The resulting fitness solution could not be concluded when too many alternatives were produced.

The authors in [139], proposed DBSCDS-GWO and DBSC-GWO to calculate the distance of transmission and balance the load on the network. A selection of dominator nodes is deterministically chosen by DBSCDS-GWO, and the metaheuristic method GWO is used to find the optimal location. Authors have also tried to minimize the Effective Transmission Distance (ETD) parameter, as follows:

$$\text{ETD} = \Phi \text{ or } \eta \text{ (whichever is shorter)} \quad (1)$$

Authors achieved stability and balanced the load of the network up to an extent, but GWO has a bad local searching ability, low solving precision, and low speed of convergence.

3.4 Security Enhancement in WSN

In past years, a large increment has been seen in the field of WSNs devices. These devices are connected wirelessly with each other and communicate wirelessly without any human intervention. During transmission of data or information from sensor nose to sink, security can be compromised. WSN protection is very important. Otherwise, it can lead to significant losses of property and life as well. Data or information can be modified or removed by the attackers to break the security of WSNs. Attackers try some specific attacks (Active or Passive) to alter the specifications of the network. These attacks are classified according to the layers of WSNs. Each wireless environment consists of the following four layers on which different types of attacks are imposed in [140].

To protect these layers from attacks and alterations, there's a need to follow some security protocols. So, three main factors need to describe the security of WSNs, which are:

- ❑ **Attacks in WSNs:** There are primarily two kinds of attacks: active (aggressive) attacks and passive attacks. Active attacks are DOS attack, physical attack, routing attack, node replication attack, node outage attack, node malfunction attack, false node, etc. Passive attacks are traffic monitoring, traffic analysis, and eavesdropping.

- ❑ **Security goals of WSNs:** WSN security goals are divided into two parts, i.e., primary and secondary goals. Primary goals are mainly confidentiality, authentication, integrity, and availability. Secondary goals are secure localization, and resilience to attacks.

- **Security protocols for WSNs:** Authors have identified various types of protection in [141][142]. WSN protocols based on SPIN, Line Selected Multicast(LSM)[143], Randomized Efficient and Distributed(RED) [142], CRS-A[144], Logical Key Hierarchy based Model(LKHM)[145], Localized Encryption and Authentication Protocol(LEAP), Certificate-less Effective Key Management(CL-EKM)[146], Low-Storage Clone Detection(LSCD)[147].

IoT is the upcoming era of communication. According to recent reports, there will be around 9.5 billion connected devices in 2020. IoT applications are rapidly increasing all over the world but are majorly used in North America, Western Europe, and China. To expand the digital economy, IoT is the major upcoming market, according to data declares in reports. Recent and existing applications of IoT are highly promising and more efficient for the users, but high-level privacy and authentication for these applications is required. Existing solutions using different technologies like machine learning, fog computing, edge computing, and Blockchain are implemented for high-level security and privacy in IoT.

3.5 Problem Statement

Due to the limited-power wireless nodes with small CPU and memory capacities, researchers were intensely interested to overcome the challenges in the WSNs field. Researchers have used conventional clustering algorithms to achieve energy-efficient WSN by taking specific distance as a parameter. The development in this sector will be the result of current technology advances and application requirements. So, by changing the distance and energy parameters, we can improve the efficiency of WSNs.

The choice of node distribution area is one of the routing problems and obstacles in WSNs that might have an impact on the routing process. To overcome routing problems, we must take care of the factors like data transmission atmosphere, low energy consumption model, fault tolerant network, QoS, and scalability of network. Load balancing to enhance the lifetime of CHs is another factor that effects the performance of WSNs. Because the nodes are situated in remote areas and rely on inadequate power source for communication, unbalancing is the biggest problem with the network. Therefore, in order to support this, the network should not be impacted by node failures caused by power supply issues. Later we need to do the research on hybridization techniques by selecting the desirable features for each to improve the efficiency of WSNs.

3.6 Optimization Algorithms

The goal of the optimization process is to discover the best solution to a given issue. The selection of an appropriate algorithm is critical for achieving this goal. However, certain issues are complicated, and finding all feasible solutions is challenging. Several meta-heuristic algorithms have been created in the literature to simulate the biological behavior of animal or insect groups by creating deterministic or random rules to be used in addressing various optimization issues.

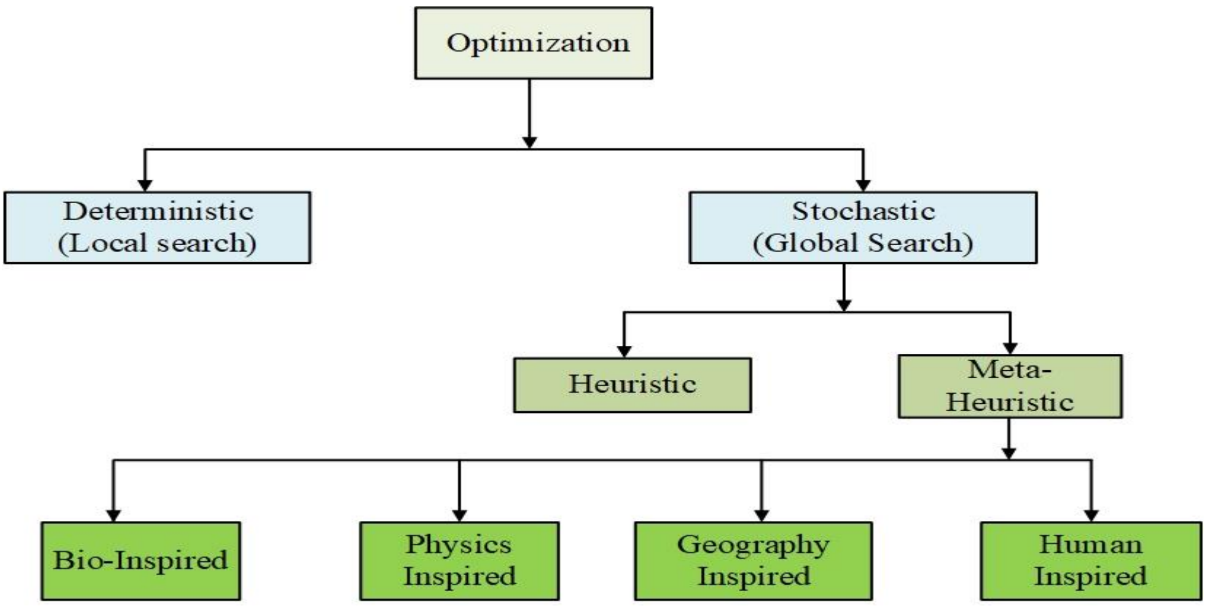


Fig. 3.1. Classification of Optimization Algorithms

Optimization algorithms helped in improving the network lifetime, stability period, throughput, number of dead nodes per iteration, and residual energy of the network. Sometimes, these bio-inspired algorithms evaluate incorrect solutions for some real-time applications. Convergence speed, multiple objective problems, dynamic problems, and local optima convergence are hot research problems nowadays. Both deterministic and stochastic optimization techniques exist in optimization algorithms as shown in Fig. 3.1.

In contrast to stochastic approaches, which merely offer a probability-based assurance, deterministic methods give us a theoretical guarantee that we will attain the global minimum or at least a local minimum. Nonetheless, stochastic approaches are quicker than deterministic ones. Moreover, stochastic approaches are appropriate for the construction of black boxes and poorly behaved functions. The deterministic technique, in contrast to stochastic methods, primarily focuses on the theoretical presumptions regarding the formulation of the issue as well as on its principal reason.

Moreover, heuristic and meta-heuristic algorithms are included in the classification of stochastic techniques. When it is challenging to identify an optimal solution, both types of algorithms are utilized to speed up the search for a global optimum. Problem-dependent algorithms include heuristics. They are extremely likely to become trapped at local optimums as a result of their adaptability to the situation and hungry character, which prevents them from reaching the global optimum. Researchers have suggested the use of creating and optimizing a multi-objective function with a suitable mathematical function-based optimizer or hybridization technique to solve challenging, dynamic and multi-objective problems in WSN.

Meta-heuristic or intelligent optimization algorithms are sometimes known as nature-inspired algorithms. The natural environment serves as inspiration for these algorithms. There are four types of nature-inspired/meta-heuristic algorithms: bio-inspired, physics-inspired, geography-inspired, and human-inspired. Biological systems are the source of the great majority of nature-inspired algorithms. As a result, bio-inspired algorithms (biology-inspired) comprise a large portion of nature-inspired algorithms.

3.7 Bio-inspired Optimization Algorithms

Recent advances in bio-inspired optimization algorithms seek to solve the issues of classical optimization methods, which are potentially providing solutions to tackle complicated optimization problems. To achieve the best results, it employs techniques that are influenced by natural selection. Selection, cross-over, and mutation are examples of such techniques. Researchers have also utilized hybridized GA with different bio-inspired algorithms. Hybrid techniques are more accurate and requires less time to complete. The techniques work well in terms of accuracy and time complexity due to the rise in the population vector size.

The main issues in this domain are primarily the optimal routing path in each cycle, data maximization with increased network lifespan, and contact distance reduction. By carefully selecting the design parameters and employing superior hybrid methods, efficiency and forecast accuracy might be improved. Under the Swarm-based techniques, four different and unique techniques are listed as shown in Fig. 3.2.

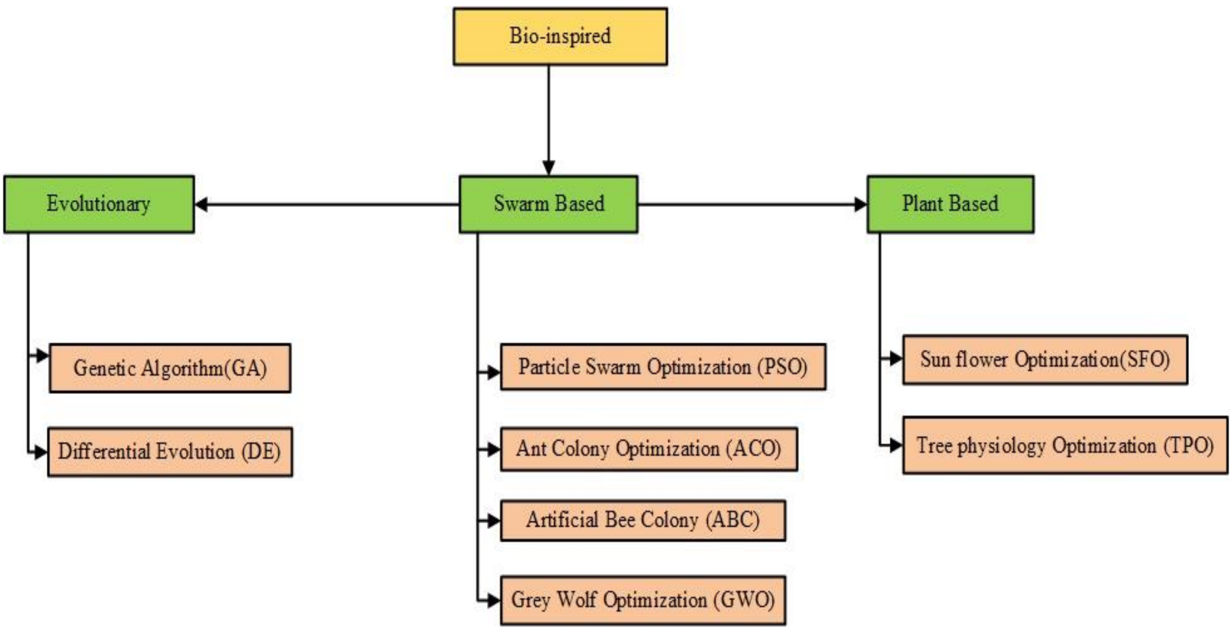


Fig. 3.2. Taxonomy of Bio-inspired Optimization Techniques

During comparative analysis, these techniques performed well in terms of increasing network longevity and energy efficiency but failed spectacularly in terms of improving QoS and security. Below explained are some important algorithms selected from a large number of nature-inspired algorithms. Based on the merits of these algorithms and their linkages to self-organization, the following algorithms play an important role in the hybridization of algorithms in WSN research.

3.7.1. Genetic Algorithm (GA)

GA was proposed by John Holland in 1960. It is an adaptive heuristic algorithm used in machine learning and artificial intelligence. This algorithm is based on natural selection and is also focused on generating optimal global solutions for optimization problems. Individual and population are two basic terms used in GA. In terms of GA, the individual is considered as a possible solution for a given problem, and a group of these possible solutions is considered as a population. Such a population of individuals is maintained within a search space. Initialization, selection, cross-over, and mutation are some important operators used by GA [31], [148], [149].

3.7.2. Differential Evolution (DE)

DE was proposed by Rainer Storm and Kenneth Price in 1997. Ever since this algorithm has been widely used in different areas like engineering science, decision sciences, material sciences, energy, etc. DE is a population-based stochastic approach in which each solution is referred to as a genome or chromosome. Each chromosome goes through mutation and recombination. DE uses terms such as target vector, donor vector, and trail vector. Only after all trail vectors have been generated is a superior solution chosen. This method also does greedy selection between the target and trail vector [8], [150], [151].

3.7.3. Particle Swarm Optimization (PSO)

J. Kennedy and R. Eberhard proposed PSO in 1995. It's a swarm intelligence approach that uses the collective behavior of birds and animals to solve optimization issues. Self-organization and division of work are two essential characteristics of swarm intelligence activity. Interactions in self-organization are carried out only based on local knowledge, with no regard for the global pattern. Positive and negative feedback, oscillations, and numerous interactions are all part of it. Tasks done concurrently by specialized persons are referred to as division of labor. The social behavior of PSO is modeled by bird flocking and fish schooling, in which each particle/bird has a position and velocity. To escape predators or find optimal environmental conditions, these particles may alter their location by changing their velocity. The velocity of the particles may be changed by modifying the particle's/or bird's group's flying experience [56], [152], [153].

3.7.4. Ant Colony Optimization (ACO)

In 1992, Marco Dorrigo suggested ACO. Ant colonies are socially complicated, with the queen as the leader and the workers hunting for food and defending the colony. Ant colonies refer to not only the physical structure in which ants reside but also the social principles by which they organize themselves and the job they accomplish. Ants have been able to use their surroundings because of their cooperation and division of work, as well as their well-developed communication systems. Ants are attracted to the pheromone trails made by other ants. If there is any obstacle on the way then ants quickly find the shortest diversion. Ant colony optimization is an optimization method that takes inspiration from the bio-semiotic communication between ants. Each constructs a solution using a stochastic greedy method using a combination of a heuristic function and pheromone trail following. ACO is related to the class of algorithms known as swarm optimization used to solve the graph search problems [27], [154], [155].

3.7.5. Artificial Bee Colony (ABC)

Dervish Karaboga proposed ABC in 2005, drawing influence from honey bees. The employed bee phase, spectator bee phase, and scout bee phase are the three periods in which bee movement is recorded in ABC. In the employed phase, the number of employed bees is equal to the number of food sources. During the employed bee phase, all solutions have the possibility of developing a novel solution. A partner is chosen at random, but the partner and the present solution should not be the same. As in the onlooker phase, the probability value of all solutions is determined before the onlooker phase. A solution with a greater fitness value has a better chance of succeeding. A fitter solution may undergo the onlooker bee phase more than once. In the scout bee phase, we have to find an abandoned solution based on the value of the limit. If some iterations exceed the defined limit, the process enters into the scouting phase and generates a new solution randomly [69], [156], [157].

3.7.6. Gray Wolf Optimization (GWO)

Mirjaliali Mohammad and Lewis presented GWO as a meta-heuristic method in 2014. The social hierarchy and hunting methods of grey wolves inspired this algorithm. These wolves lived in well-structured packs, with several wolves ranging from 5 to 12. The members of the pack are divided into four categories i.e., α-wolves, β-wolves, δ-wolves, and ω-wolves. Alpha wolves are the leaders of the pack and the rest of the pack follow alpha. Alpha wolves are in charge of making decisions regarding hunting, sleeping, and waking up times, among other things. There are beta wolves who are the greatest contenders to be alpha at the second level. Delta wolves are present in the pack to supply food and to protect the pack in times of danger. Omega wolves (Scouts, Elders, and Caretakers) are at the bottom of the food chain, serving as scapegoats and the last to eat. The following are the main phases in the GWO hunting process: 1) Searching for prey 2) Tracking, pursuing, and approaching the prey 3) Encircling and tormenting the prey till it finally comes to a halt; 4) Taking on the prey [29], [158]–[160].

3.8 Selective Bio-inspired Hybrid Optimization Algorithms

The present meta-heuristic techniques have several drawbacks, including sluggish convergence and limited accuracy. Scholars have progressively turned their attention to the swarm intelligence algorithm in recent years. Swarm intelligence algorithms are widely used because of their simplicity, adaptability, non-derivation mechanism, and avoidance of local optimality. The features and trends of scientific growth are reflected in the rapid development of swarm intelligence algorithms. In this paper, we're looking at some new hybridization approaches to nature-inspired algorithms to make the algorithm more resilient and enhance simulation analysis and outcomes statistics.

3.8.1. Hybrid GA-DE Algorithm

GA includes solutions regarding non-convex and nonlinear problems. As we know, different operators like initialization, selection, and cross-over are used by GA. In hybridization of GA-DE, mutation operation is performed by DE. DE also solves non-differential and non-continuous real-world problems. So, hybridization of GA-DE would be able to provide better global optimal solutions [32], [161].

Design procedure for GA-DE algorithm:

- ❑ First SN control variables are selected like genes.
- ❑ Initialize the population of sensor nodes.
- ❑ Using the localization function, calculate the fitness of SNs.
- ❑ Use the roulette wheel selection method for mating.
- ❑ GA performs cross-over operations.
- ❑ Mutation operation is performed by DE.
- ❑ Select a new population for the upcoming generation.
- ❑ Repeat steps four, five, six, and seven.
- ❑ Print estimate of location.

In the performance research of a hybrid GADE localization algorithm with localization function, the hybrid technique's precision and time complexity in the context of varying population vector sizes with localization function are reported. In comparison, when the size of the population vector grows larger, the accuracy improves and the time complexity performance improves. Also, when the hybrid GADE localization algorithm employs the average localization function instead of the basic localization function, it outperforms the better competition in terms of temporal complexity and accuracy.

3.8.2. Hybrid GA-PSO Algorithm

The fundamental goal of hybrid GA-PSO is to enhance CH selection and routing between deployed nodes and the base station. There are two steps to the proposed method. In the first phase, the PSO algorithm holds passed population and fittest individuals. In the second phase, these fittest individuals are operated by the GA operators who are selection, cross-over, and mutation. Hybridization of GA and PSO combines the merits of both algorithms, which provide us better convergence rate and avoid the problem of local optima [33], [162]. The major contribution of both algorithms has been shown in Table 3.1.

Table 3.1. Combination of GA-PSO technique in WSNs

Algorithms	GA	PSO
Operators used	Selection, cross-over, mutation	Inertia, cognitive, social
Ability to search global optima	High	Low
Implementation	Hard	Simple
Trapped on local optimum	Sometime	Often
Computer efficiency	Low efficient	Highly efficient

This method combines the benefits of both algorithms, such as PSO's rapid convergence rate and GA's problem of trapping in local optima. The primary purpose of this PSO-GA strategy is to steadily raise the number of decent people across generations.

Design procedure for GA-PSO algorithm:

- ❑ Initialization
- ❑ Generation of the initial population
- ❑ Selection
- ❑ Cross-over
- ❑ Mutation
- ❑ Growth
- ❑ Generation of a new population
- ❑ Repeat until no. of generation (Ng) evaluates.

The hierarchical sensor network model is used in the PSO, GA, and PSO-GA approaches for small and large size networks. In a hierarchical WSN paradigm, each cluster contains one base station and one relay node. In this WSN configuration, the relay node serves as the CH. The most significant assumption is that base stations offer routing pathways and that each relay node's average data volume is known. Each relay node's leftover energy is replenished at the end of each generation, and current energy is utilized to determine the next routing path.

When comparing the shortest path method, PSO approach, GA approach, and hybrid PSO-GA approach for large-scale networks, we observed that the hybrid PSO-GA strategy has the best network lifespan and packet delivery ratio.

3.8.3. Hybrid ACO-PSO Algorithm

The main aim of ACO-PSO hybridization is to improve inter-cluster data aggregation in WSNs. This proposed technique also improves the network's lifetime over many optimization techniques. In this approach, ACO results in local updates, and PSO gives a better outcome for global updates. The combination of ACO-PSO enhances the durability and performance by 6% over previously used optimization techniques like an ant colony, cuckoo search, flower pollination, etc.

Design procedure for ACO-PSO algorithm:

- ❑ Step 1. Initialize the number of wireless sensors.
- ❑ Step 2. Calculate the energy level for each SN.
  If E>0, there will be a selection of CH otherwise, go to Step 2 again. After CH selection, implement the ACO-PSO algorithm to find a new path.
- ❑ Step 3. Calculate the energy dissipation for each SN.
  If a dead node is found, count dead nodes else, Step 2.
  For different parameters, evaluate the performance of the network.

Hybrid ACO-PSO-based data aggregation is used to increase the inter-cluster data aggregation. Extensive investigation shows that the suggested approach significantly increases network lifespan when compared to previous strategies. It divides the sensor network into several pieces, referred to as clusters, with CHs chosen for each cluster. Then, using short-distance connections, tree-based data aggregation is used to acquire sensory data directly from CHs. The use of compressive sensing decreases the size of the packets that are sent across the sensor network.

The ACO-PSO algorithm determines the shortest path between the sink and CHs. For simulation, the MATLAB simulation tool is generally utilized by researchers. It helps in comparing the proposed approach's performance to that of existing technology, GSTEB, in terms of stability period, network lifespan, residual energy (average remaining energy), and throughput.

3.8.4. Hybrid PSO-GWO Algorithm

The aim to hybridize PSO and GWO is to obtain more optimal results with less no. of iterations. This approach successfully merged the powerful merits of both algorithms to get better efficiency. Simplicity, fast convergence speed, and high exploitation ability are some of the advantages of this algorithm. When these merits of PSO collaborated with GWO of high exploration ability, it enables higher stability and shows better performance with more optimal solutions [35], [36], [163].

Design procedure for PSO-GWO algorithm:

- Set the overall population and the A, C, and a value to their defaults.
- Create people for the population.
- Find the fitness value of each individual.
- Calculate the value of α, β, and δ by shortening the order according to size.
- Calculate nonlinear controlled parameters and update the value of A and C.
- Detect the location of individuals and again calculate fitness values.
- Update the values of α, β, and δ.

The PSO approach has been utilized to address almost any real-world problem. However, there must be a mechanism to reduce the chances of the PSO algorithm catching itself at a local minimum. Recommended techniques to reduce the likelihood of falling into a local minimum have introduced the GWO algorithm to support the PSO algorithm.

To avoid risks, the GWO algorithm's exploration ability is employed to send certain particles to sites that are somewhat enhanced by the GWO method rather than random positions. Since the GWO method is used in addition to the PSO algorithm, the running duration is also increased. The PSO-GWO algorithm incorporates nonlinear control parameters. Other algorithms have inadequate nonlinearity in their control parameters, resulting in a lack of balance between local and global search abilities and an easy fall into local optimum throughout the search phase.

3.9 Comparative Analysis based on Recent Literature

A comparison of the various methodologies and criteria utilized in recent articles to choose the CH has been conducted. Each of the measures has been explained in terms of percentage, utilization, and progress [164]. These measures are packet loss, network lifetime, energy, throughput, delay, and overhead. Packet loss occurs when one or more data packets fail to arrive at their intended destination. Network lifetime is based on the number of alive nodes, connectivity, and sensor coverage. The energy of WSNs is evaluated based on the packet received at the destination. Throughput is the actual amount of information that is efficiently sent/received via a communication channel. Delay is the time taken by the packets from SNs to sink and is proportional to the no. of hopes. Overhead is the total amount of energy consumed to transmit the data for a given time.

Table 3.2 compares several approaches offered by notable scholars in terms of parameters employed, tools used, and difficulties handled.

Table 3.2. Comparative analysis of various conclusions based on recent literature

Reference	Parameters Used	Tool Used	N/w Life	Energy Efficiency	QoS	Security	Results
Thenmozhi et al. [165]	Residual energy, Node's capability Assembly compactness, Node's gradation	MATLAB	√	√	√	×	The overall delay is reduced by 23%. The rate of packet loss is reduced by 11%. Residual energy improved by 38%.
Jia et al. [166]	Area Coverage, Life rotation, Dynamic nodes, Average remaining energy.	MATLAB	√	√	×	×	In comparison to LEACH and DEEC, network lifespan increased by 50% and 30%, respectively. Clustering overhead was reduced by 42%.
Aggarwal et al. [167]	Remoteness to the sink, Enduring energy, Sensor Node concentration	MATLAB	√	√	×	×	Network lifetime increased by 30%. In comparison to LEACH and EAUCF, Prolonged Energy increased by 155.18% and 35.75% respectively.
Neamatollah et al. [168]	Residual Energy, Gradation of the sensor node, Distance SNs to BS.	MATLAB	√	√	×	×	The network's lifespan has increased by 28%. The overhead of clustering was decreased by 57%. A 13 % reduction in energy utilization.
Mehra et al. [169]	Residue Power, Base station's remoteness, Concentration of the SNs.	MATLAB	√	×	×	×	In comparison to LEACH, network lifespan rose by 15%, 11.38 % with BCSA, and 8.1 % with CAFL. Energy conservation by 79%.
Jeong et al. [170]	Concentration, Centrality, Overhead, Average delay	MATLAB	√	√	×	×	In comparison to LEACH, there is a 42.7% increase in power and local distance.
Krishna et al. [171]	Network lifetime, Throughput, Distance between SN to CH, No. of neighboring nodes.	MATLAB	√	√	×	×	Average left overpowers and alive nodes improved by 62%. Overall, this is a 45% improvement over LEACH.
Azad et al. [172]	Remaining energy, the route followed between sensor nodes and sink. No. of neighbor nodes.	MATLAB	√	√	√	×	TOPSIS has a 151.2% percent longer network lifespan than LEACH. Overall, 40% better than LEACH.
Behra et al. [11]	Network's coverage, Total number of sensor nodes, Entire network energy consumption, Energy degeneracy.	MATLAB	√	√	√	×	Packet loss rate reduces by 8%. Throughput has increased by 60%, lifetime has increased by 63%, and residual energy has increased by 61%.
Tamizharasi et al. [173]	Usual Enduring Energy Number of the active nodes, Entire nominated cluster head.	NS2	√	√	×	×	In comparison to 19% for LEACH, 5 percent of the increase in energy utilization. Increase in the number of living nodes with a longer lifespan.

During comparative analysis, researchers implemented different deterministic and probabilistic approaches. These techniques performed well in terms of increasing network longevity and energy efficiency but failed spectacularly in terms of improving QoS and security. Prioritizing and allocating time slots to the nodes with gathered information takes into account the packet loss percentage and delay. The network efficiency in WSN is improved by this approach because it avoids utilizing pointless retransmissions and delays. Simulation findings demonstrate that the suggested technique boosts the packet delivery ratio and remaining energy while decreasing delay and overhead. In Fig.3.3, researchers tried to reduce the packet loss and overhead. Also, they have made an effort to enhance the lifetime of the network i.e.; it depends on energy consumption.

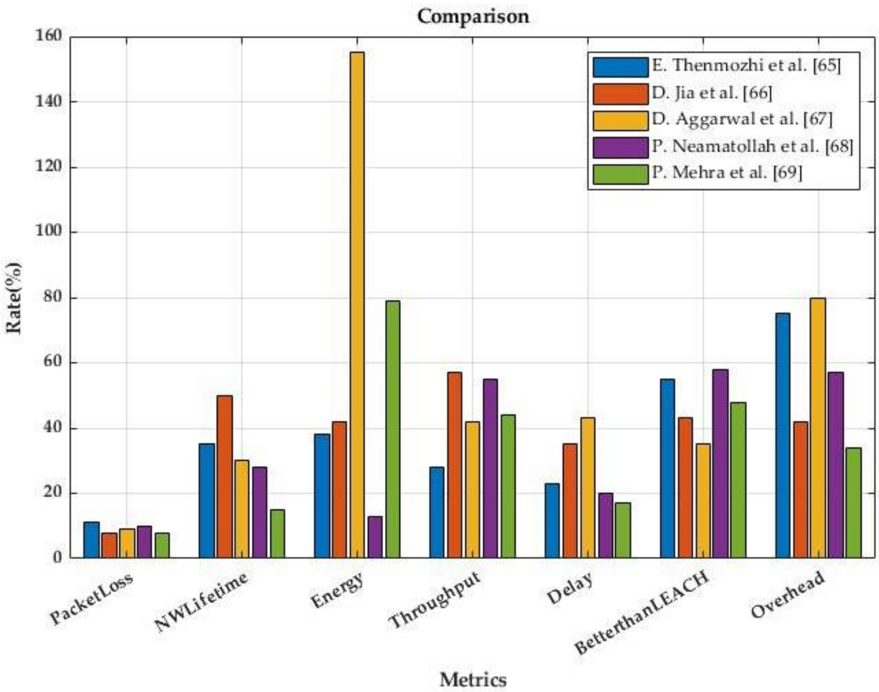


Fig. 3.3. Comparison of the various methodologies and performance metrics

A comparison of the various methodologies and performance metrics is shown in Fig. 3.3. Thenmozhi et al. [165] reduced the packet loss by 11%. In comparison to LEACH and DEEC, network lifespan increased by 50% and 30%, respectively to Jia et al. [166]. In comparison to LEACH and EAUCF, Prolonged Energy increased by 155.18% and 35.75% respectively to Aggarwal et al. [167]. Neamatollah et al. [168], and Mehra et al. [169] also improved the overall performance with a high reduction in delay and overhead, respectively.

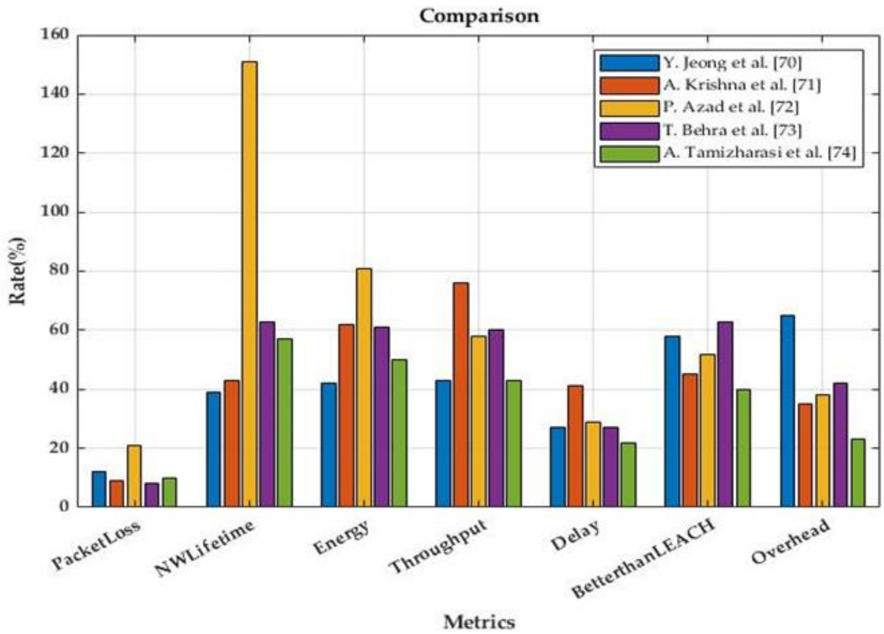


Fig. 3.4. WSNs performance comparison based on recently reviewed articles

In Fig. 3.4 researchers tried to improve the lifespan of the network by taking energy and throughput into their account. The authors also compared their results with the existing algorithms. Jeong et al. [170] and Krishna et al. [171] improved in energy and overall performance by 43% and 45%, respectively, as compared to LEACH. Azad et al. [172] improved the maximum energy residual and network longevity, according to the results. Behra et al. [11], and Tamizharasi et al. [173] also boosted the maximum energy and network longevity. The algorithm improved in extending the life of the network and preserving leftover energy.

3.10 Comparative Analysis with Hybrid GA-PSO Technique

3.10.1 Operational Structure and Flow Process for Hybrid GA-PSO

- The different phases outline how the GAPSO-H functions exceptional:

Step 1: The sink is originally situated in one of the network's corners, and the diverse nodes are distributed randomly across the system.

Step 2: The preprocessing stage, which follows network establishment, is the next stage. GA is used to pick the CH in this process. During the cluster formation, the CH is chosen, and during the saturation phase, the cluster members communicate the CH.

Step 3: Replacing the BS using the PSO technique is the next operation. It will assist the sink in determining where it need to move in order to gather data.

Step 4: According to the energy model, the nodes use their energy.

Step 5: The technique is then put into practice in MATLAB, and the simulation results are collected.

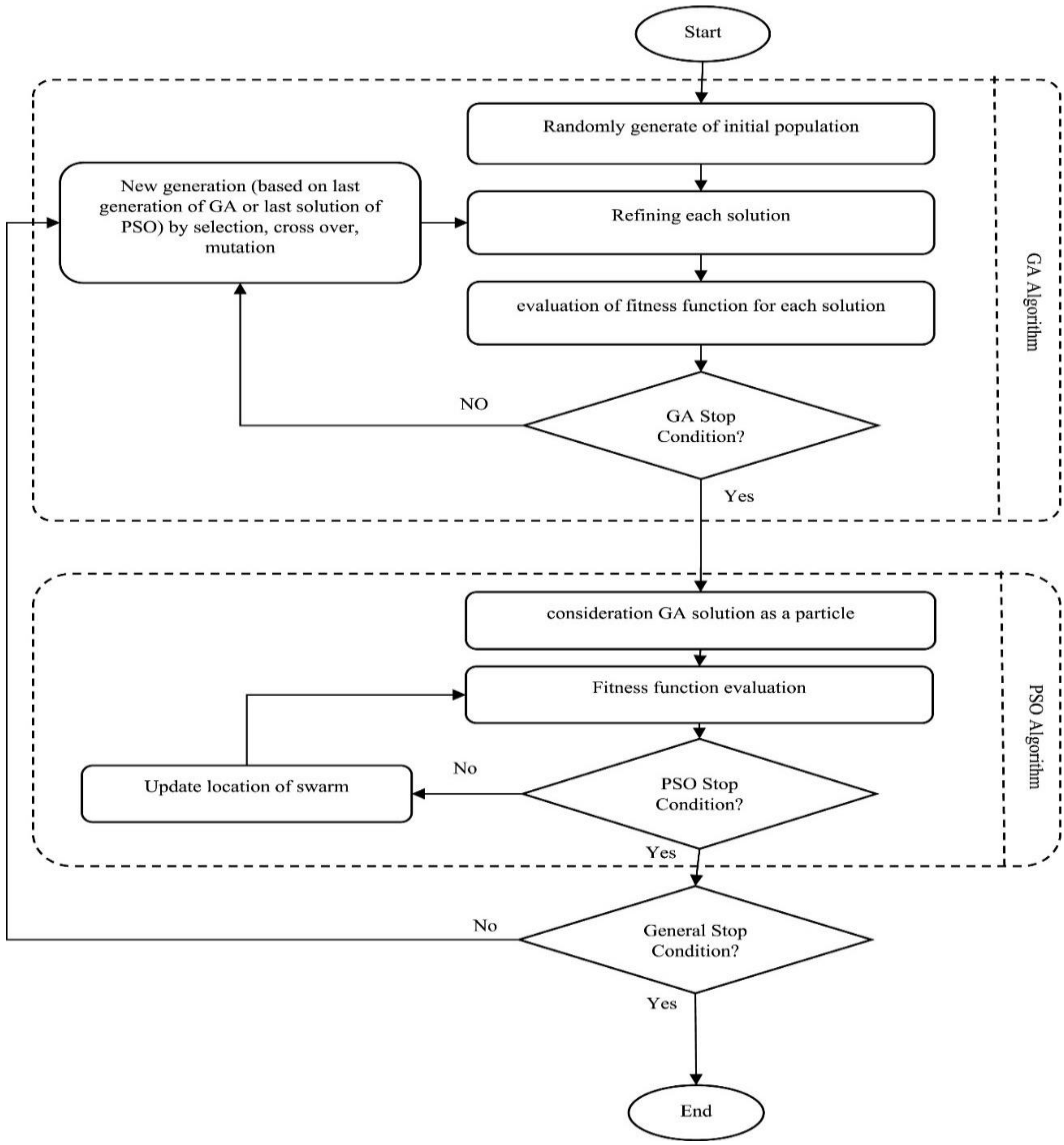


Fig. 3.5. Flow process of GA-PSO model

- The amalgamation of GA and PSO as shown in Fig. 3.5 has been used to resolve different real-life problems. The flow-process for GA-PSO is as follows:

Step I: Random generation of the initial population of chromosomes.

Step II: The cost of implementing every solution from step 1 is calculated as the survival of each chromosome.

Step III: The GA algorithm's process will halt when the stop condition is met and continue in a different one.

Step IV: By using selection, crossover, and mutation processes, produce an entirely novel population.

Step V: Go to step II.

Behra et al. [11], and Tamizharasi et al. [173] also boosted the maximum energy and network longevity. The algorithm improved in extending the life of the network and preserving leftover energy.

3.10.2 Performance Matrices

The intended GAPSO-H has been evaluated using the five-quality metrics that are detailed below:

- **Stability Period:** It may be described as the quantity of rounds completed before any node's energy reserves are totally depleted. As a result, the node is considered to be dead. In cases where data leaking cannot be tolerated, this value might be very important. Applications with such a limited tolerance for information loss is guaranteed with larger stability period values.
- **Network Longevity:** It may be described as the rounds completed when there are no nodes surviving in the network that can interact. For applications where security management is a continual activity, it is essential. Applications like farming, flood detection, etc. are a few instances where network lifespan is quite important.

- **Number of dead nodes:** The rate at which the nodes continue to lose energy until they are completely dead is tracked in to track the behavior of the protocol when it experiences energy consumption during packet forwarding.
- **Throughput:** Throughput is the term used to describe the quantity of data packets that are sent after a successful transmission. It aids in describing each node's QoS state. In the event that the QoS is subpar for any routing protocol, the stability period and network lifetime are worthless.
- **Network's remaining energy:** At varying numbers of rounds, the network's remaining energy is shown. It displays the network's task scheduling. The total energy of all nodes distributed in the network is essentially the network energy.

3.10.3 Results and Comparative Analysis

Comparative investigation shows that, in terms of several performance criteria, GAPSO-H has surpassed other cutting-edge algorithms. In terms of stability period, HND, last node dead (also known as "Network Lifetime"), and throughput, the percentage improvement made by the GAPSO-H is shown in Table 3.3.

Table 3.3. Enhancement in Percentage (%) by GAPSO-H compared to other protocols

Algorithms	Stability Period	HND	LND or Network Lifetime	Throughput
DCH-GA [174]	72.6	79.0	69.6	78.6
GABEEC [175]	44.8	42.2	35.7	39.6
PSOBS [176]	39.5	31.8	19.0	63.0
GADA- LEACH [177]	44.3	65.2	46.1	56.6
PSO-UFC [178]	14.9	26.1	27.1	13.5
PSOECSM [179]	5.5	15.7	12.0	24.3

Comparing GAPSO-H to DCH-GA, GABEEC, PSOBS, GADA-LEACH, PSO-UFC, and PSOECSM, respectively, in Fig. 3.6 and Fig. 3.7, GAPSO-H improves the stability period by 72.6%, 44.8%, 39.5%, 44.3%, 14.9%, and 5.5% [167–172].

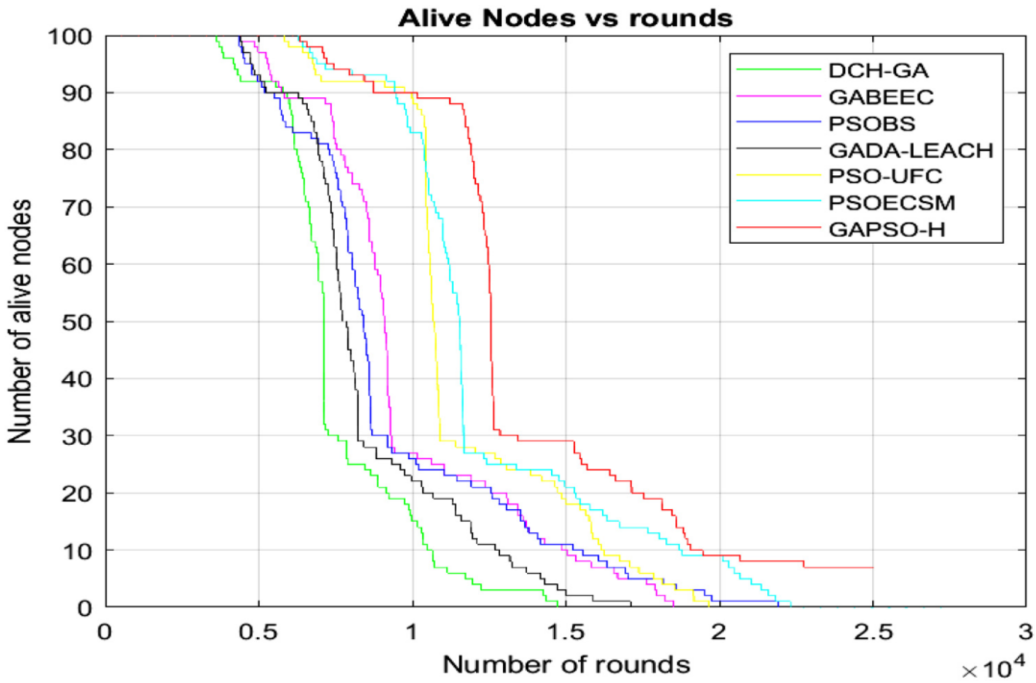


Fig. 3.6. Comparison of alive nodes vs rounds of GAPSO- H with other protocols

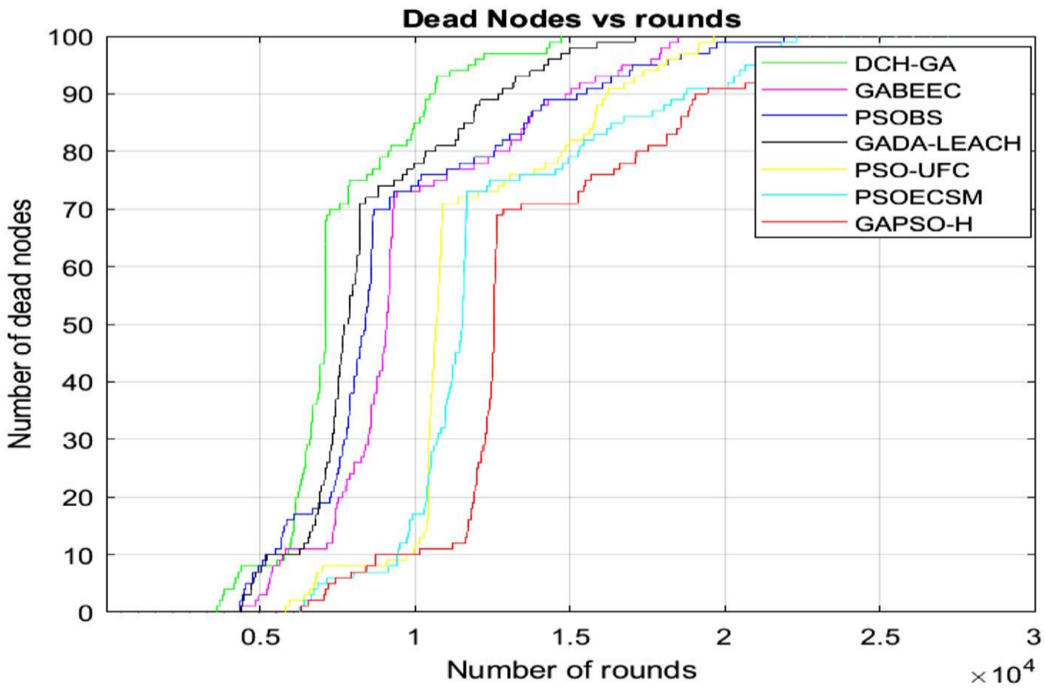


Fig. 3.7. Comparison of dead nodes vs rounds of GAPSO- H with other protocols

Fig. 3.8 and Fig. 3.9 show the network's remaining energy and packets sent to BS of GAPSO-H with DCH-GA, GABEEC, PSOBS and GADA-LEACH, PSO-UFC, and PSOECSM [167–172].

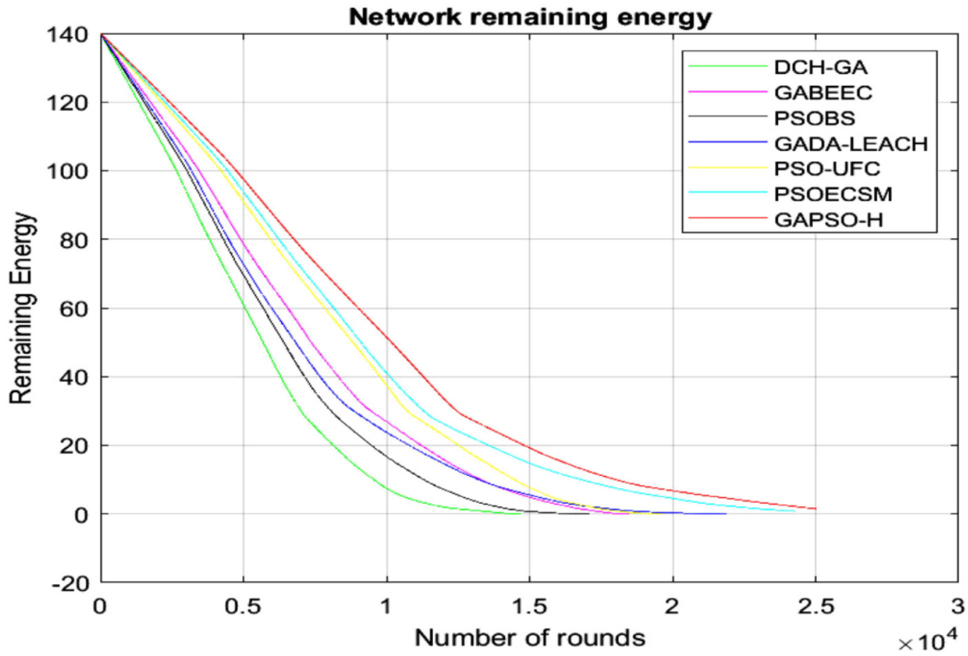


Fig. 3.8. Network's remaining energy of GAPSO-H with other protocols

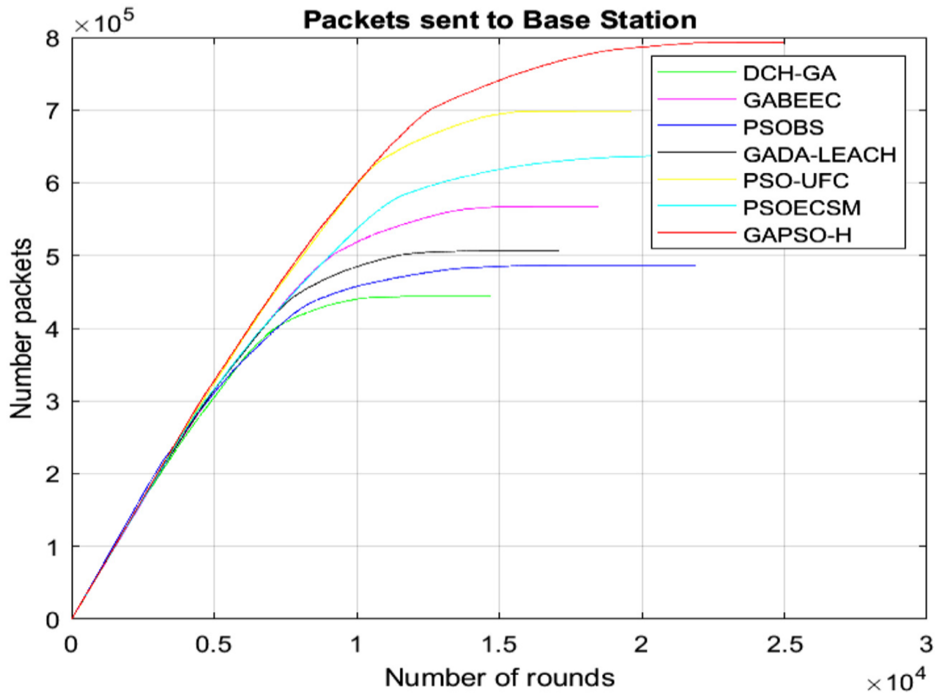


Fig. 3.9. Throughput comparison of GAPSO-H with other protocols

The network's lifetime and overall performance comparison of GAPSO-H with DCH-GA, GABEEC, PSOBS and GADA-LEACH, PSO-UFC, and PSOECSM [167–172] have been done as shown in Fig. 3.10 and Fig. 3.11.

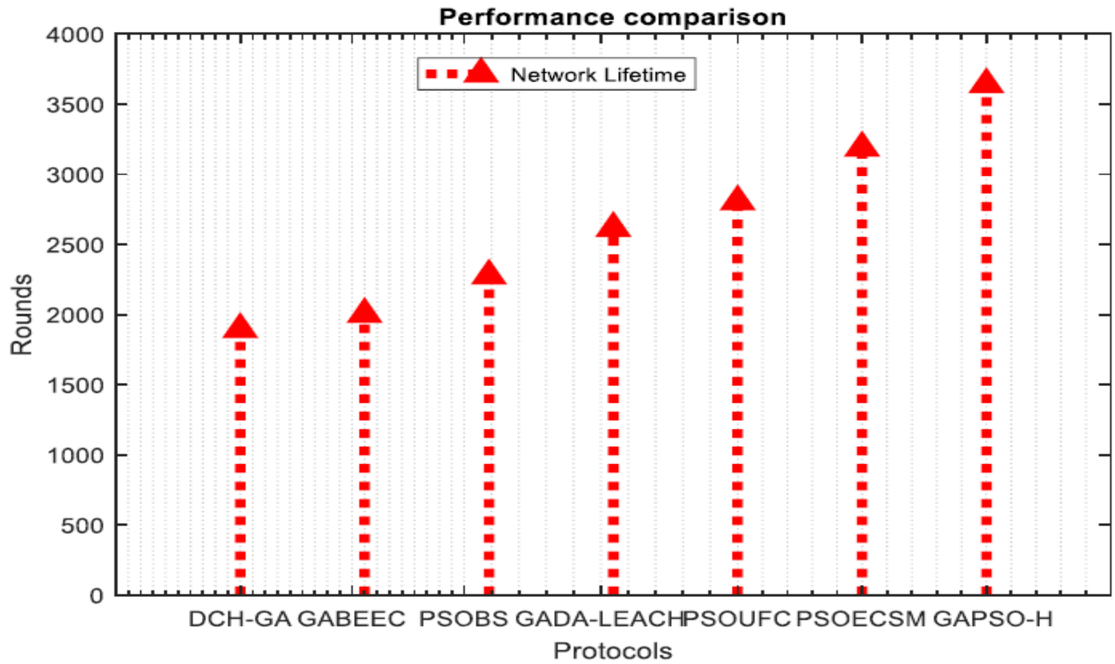


Fig. 3.10. Performance comparison of GAPSO-H with other protocols

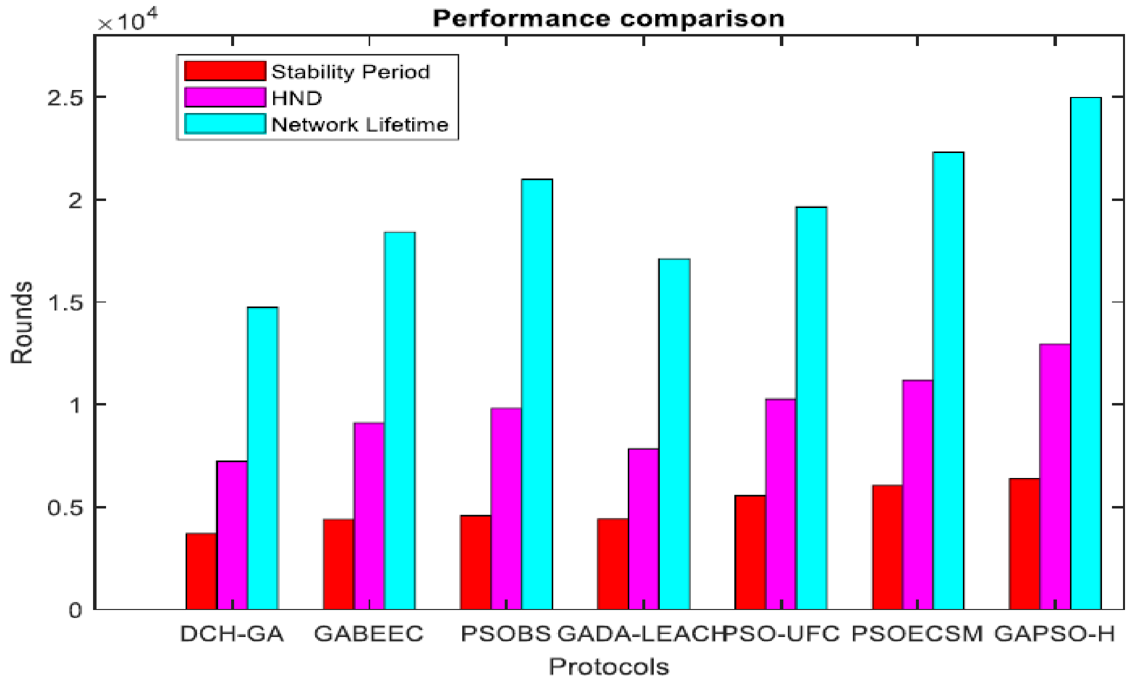


Fig. 3.11. Overall performance comparison of GAPSO-H with other protocols

3.11 Comparative Analysis with Hybrid PSO-GWO Technique

There are some drawbacks to the GWO algorithm, including sluggish convergence and limited accuracy. GWO and PSO were used to address these issues. The algorithm's global search and local search capabilities are balanced, and the method's convergence speed is increased using a nonlinear control parameter. PSO, which updates each grey wolf's position data using the best values of the individual and the wolf pack, is also implemented at this time. The suggested approach prevents the algorithm from reaching a local optimum and maintains the individual's best position information as shown in Fig. 3.12.

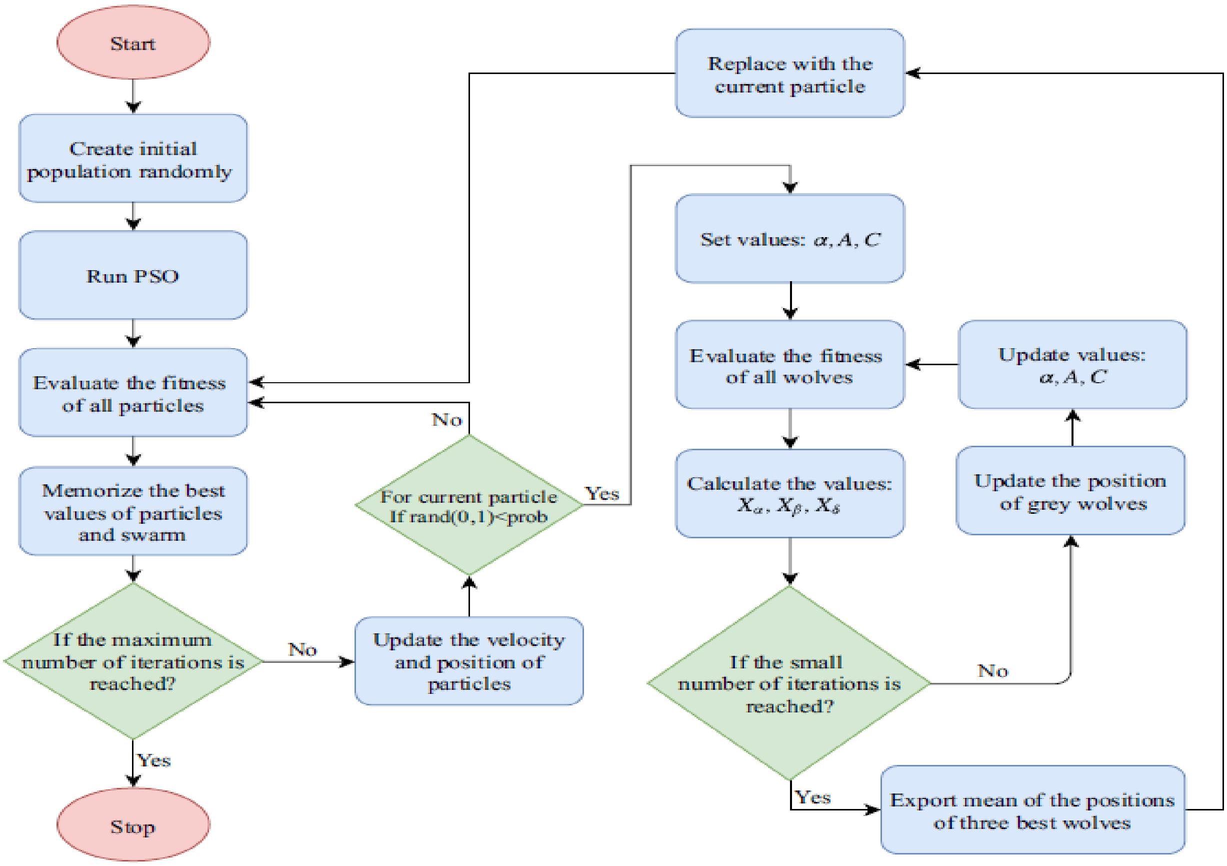


Fig. 3.12. Flowchart of HPSGWO

Step 1: Initialize the A, C, and a values while setting the population size to N and dimension d.

Step 2: Create population entities, then calculate the individual fitness value.

Step 3: Sort the fitness values in ascending order by size, then choose the top three fitness values that correspond to the person as α, β, δ. Information about the associated location is Xα, Xβ, Xδ.

Step 4: Update the values of A and C after calculating the nonlinear control parameter.

Step 5: Change each individual's location before recalculating their fitness scores and updating the values for α, β, δ.

Step 6: Analyze if t reaches the Tmax value; if it does, the result is the fitness value α, which is the optimal answer.

Otherwise, proceed to Step 3.

The continuous statistics on the nodes and route are delivered by hybrid PSO-GWO. Using the details from the path database, it determines the legitimate path and the optimum route. As a result, Hybrid PSO-GWO functions as an energy-efficient WSN for safe data transfer.

Fig. 3.13 shows that for network 1, hybrid PSO-GWO uses 9% less energy than PSO [56] and 10% less than GLBCA [134] and GA [180], whereas for network 2, hybrid PSO-GWO uses 13% less energy than PSO, GA, and approximately 15% less than GLBCA.

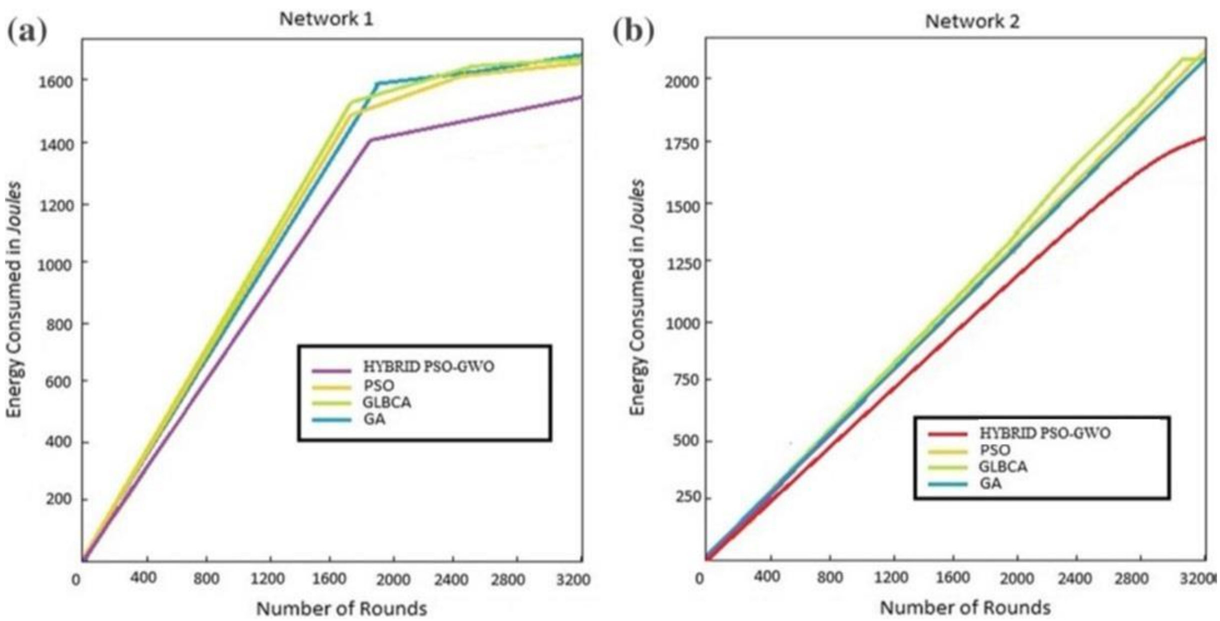


Fig. 3.13. Energy consumption for Network 1 and Network 2

Figure 3.14 shows that for 600 nodes, the lifetime of the proposed algorithm is increased by 18% and 41% compared to PSO and GLBCA, nearly 38% compared to GA, and 63% compared to LDC. For 700 nodes, the lifetime is increased by 16% compared to PSO [56], 48% compared to GLBCA [134], 57% compared to GA [180], and 75% compared to LDC [181].

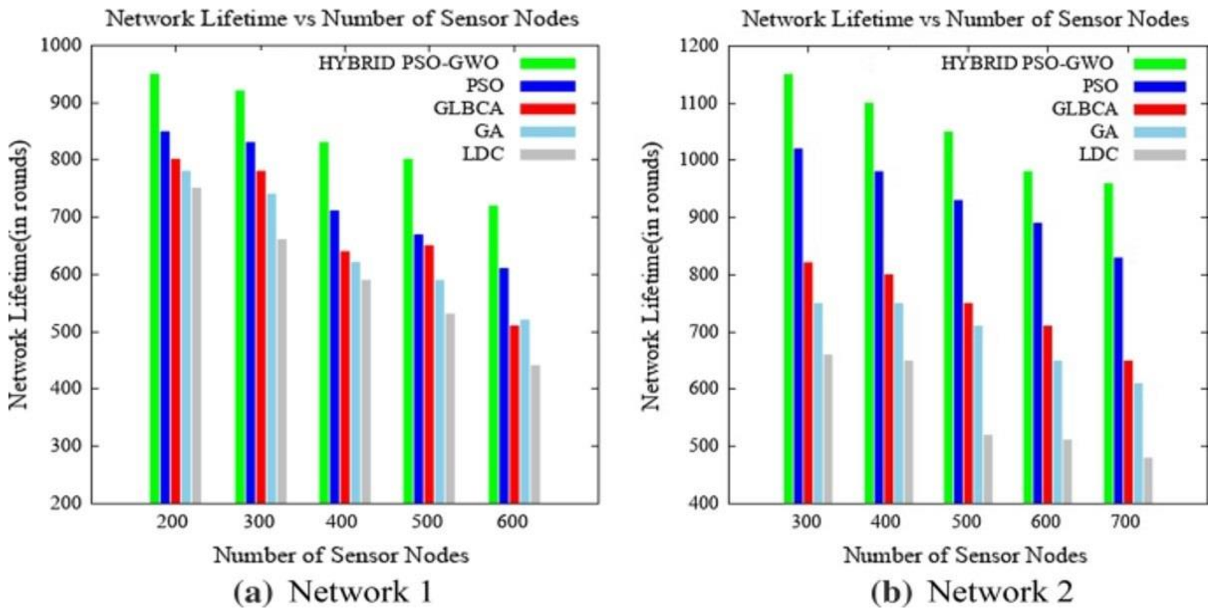


Fig. 3.14. Network lifetime for Network 1 and Network 2

3.12 Conclusion

Different optimization algorithms viz. Genetic Algorithm (GA) [25], Particle Swarm Optimization (PSO) [75], Ant Colony Optimization (ACO) [34], Biogeography-Based Optimization (BBO) [70], and Gray Wolf Optimization (GWO) [28] are proposed in the literature to evaluate the performance and security of WSNs. It is widely acknowledged in the literature that integrating many complementing characteristics derived from the applications improves accuracy and performance. To provide academics with future directions, this review study examines current trends in the area of WSN security analysis. In this area, there is a lot of research being done to find a trustworthy assault detection method. It is well known that ML-based approaches offer reliable answers, but their implementation necessitates enormous training data sets and specialized technology. Large datasets have also been created in order to meet the ever-increasing need for training data. This research work makes use of hybrid algorithms that are inspired by nature in order to maximize the efficiency and security of WSNs. These algorithms improve the network lifetime while simultaneously reducing energy consumption and processing delay in order to respond to environmental factors that are helpful for the targeted situation. Yet, new malware variants are discovered every day, and keeping up with the most recent patterns effectively to recognize emerging patterns in the use of permissions is another ongoing issue.

Chapter Related Publications

- **Rajiv Yadav,** Indu Sreedevi, Daya Gupta, "Bio-Inspired Hybrid Optimization Algorithms for Energy Efficient Wireless Sensor Networks: A Comprehensive Review," *Electronics* 2022, 11, no. 10, 1545. **SCIE, (IF: 2.690),** (Accepted & Published), DOI: 10.3390/electronics11101545.
- **Rajiv Yadav,** Indu Sreedevi, Daya Gupta, "Review of Evolutionary Algorithms for Energy Efficient and Secure Wireless Sensor Networks," International Conference on Cyber Security and Digital Forensics, pp. 597-608. **Springer, (Scopus Indexed),** Singapore, 2022.DOI: 10.1007/978-981-16-3961-6_49

Chapter 4

Design and Development of an Improved Energy-efficient Nature-inspired Algorithm in WSNs

4.1 Introduction

Wireless communication and electronics advancements resulted in the creation of low-power, low-cost, multipurpose WSNs[182]. WSNs are made up of self-configured, scattered, and self-governing SNs that display corporal and ecological measures such as moisture, temperature, and sound in a given deployment region[183][184]. Because sensor node energy supply and transmission range are restricted, a suitable approach for calculating energy-efficient pathways to relay data from the SNs to the BS is required[185][186]. In a cluster-based WSN, each has a front-runner known as the CH. All of the SNs perceive statistics and send them to their associated CH, who then refer them to the BS for processing. Data aggregation is performed inside a cluster to decline the quantity of data to be shifted to BS. Second, rotating CHs helps maintain the balanced energy usage inside the network, and thus prevents individual nodes from becoming energy-starved[187]. However, choosing the right CH with the best capabilities while balancing the network's energy efficiency ratio is a precise NP-hard problem[188]. Clustering methods aim to split the network's SNs into suitable clusters and choose cluster leaders to link directly with the BS. In large-scale WSNs, the majority of SNs are located distant from the BS and are unable to communicate directly. The majority of clustering algorithms ignore how CHs link with the BS. The following issues may emerge in a highly-populated network[189]:

- ❑ There are a lot of nodes that are interacting with each other.
- ❑ A variety of data transmission paths.

- ❑ Individual nodes communicate with distant nodes, wasting energy unnecessarily.
- ❑ Every modest topological change necessitates the creation of a new routing path.

The Butterfly Optimization Algorithm (BOA), a novel optimization algorithm inspired by nature, is based on how butterflies find food. In order to address issues with global optimization, BOA imitates the feeding and mating habits of butterflies. With their sense of smell, animals can locate sources of nectar or potential mates. The three key terms of sensory modality (c), stimulus intensity (I), and power exponent (p) form the foundation of the entire idea of perceiving and processing the modality (a). The terms sensory modality and modality relate to the raw input that the sensors utilize to measure and process energy, respectively.

BOA is employed in WSNs to find the ideal path among the CHs to save energy. Fault tolerance, dependability, information accumulation, scalability, and other characteristics of routing protocols are among them[190]. IBOA selects the best CH from a set of nodes to expand the stability, merging speed, and problem of trapping in local minima. The remaining energy, distance from SNs to BS, node degree, and node centrality play an important role in CH selection[191]. Active nodes, energy usage, data packets acknowledged by the BS, and dead nodes are used to estimate the proposed methodology's act. When comparing the anticipated method to existing techniques, the planned procedure outperformed LEACH[192] and ZSEP[193] in terms of network lifespan. It was discovered that the recommended approach outperformed current methods in terms of network performance. The scenery of the goal function affects the sensory intensity of a butterfly[194]. The inspiration behind this research is to condense node energy usage during data transmission. This diminishes the energy ingesting of the SNs, allowing for more overall packet transfer to BS. Because of its searching capability, resilience, and self-adaptability, swarm intelligence is mostly employed in this study.

4.2 Related Work

D. Karaboga [156] deliberated an ABC algorithm to outspread the network’s era, an advanced energy-efficient clustering procedure implemented on ABC. Clustering approaches have successfully deployed the ABC algorithm that pretends the intelligent foraging conduct of honey bee clouds. The suggested approach's performance is equated against protocols based on LEACH and PSO, which have been investigated in a variety of steering situations. The suggested technique is compared to LEACH and PSO-based protocols. This technique not only extends the network’s lifetime but also implements a service value phenomenon by taking into account delays between indicators acknowledged from clusters. P. Kuila [56] proposed linear and nonlinear optimization problems for clustering and routing. The routing technique uses a multi-objective fitness function and an effective particle programming approach in PSO. And the clustering approach is described by taking into account the nodes' energy saving through load balancing. It also searches a search space of candidate solutions for the best solution without using gradient-like other optimization algorithms do.

T. Bhatia [195] proposed the GADA-LEACH technique that uses an evolutionary GA to improve CH assortment in the standard LEACH routing protocol in WSNs. It is more efficient than the former because it includes a larger quantity of parameters for picking better CHs. The addition of a transitional node such as a relay node shortens the remoteness between CH and sink. The simulation findings suggest that GADA-LEACH outperforms other traditional protocols of LEACH in standings of network lifespan. P Lalwani [196] proposed a firefly algorithm to outspread the life of WSNs. To demonstrate its performance, the projected method is thoroughly established under variability of circumstances and compared to different algorithms. The suggested algorithm was compared to several current algorithms in a variety of circumstances, including HF, EADC, and DHCR. When compared to existing ones, FARW is determined to be competitive/superior in the majority of scenarios in the experimental investigation.

N. Jain [197] explained a comparison of the merits of several routing protocols for WSNs. The management of the energy accessible in each SN is one of the utmost significant enterprise elements for a sensor model. In WSNs, growing network longevity is crucial. In this context, several routing methods have been created. Clustering techniques have become quite important in improving the network lifespan and hence the efficacy of the nodes in it. Clustering is a viable option for extending the life of a WSN. Authors have examined six well-known routing protocols, including TEEN, SEP, LEACH, ERP, EAMMH, and PEGASIS, for a variety of circumstances. TEEN has a higher level of stability than LEACH and SEP, according to the findings. The EAMMH and PEGASIS protocols outperformed the LEACH procedure.

M. Singh [198] proposed a clustering technique in WSNs using the firefly optimization method. The authors have shown that the concert parameters improved as a result of the simulation findings. As the energy fatigue, a node's network lifespan was found to be improved. The suggested approach's results were compared to those of the existing Energy-Aware clustering method. The packet delivery ratio improves, implying that packet loss in the network was minimized. N. Al-Aboody [158] proposed the Grey Wolf Optimizer (GWO) to create a clustering routing protocol MLHP for WSNs. The method was put to the test by measuring the energy efficiency, longevity, and stability of a network. The suggested approach outperformed the benchmark algorithm in terms of stability and network lifespan, according to simulation results and system assessment. In terms of network longevity, MLHP outperformed LEACH by 500 times.

T. Shankar [58] purposed a combination of HSA and PSO algorithms developed for energy-efficient CH selection to run a global search with quicker conjunction. The suggested technique was found to have a higher HSA search proficiency and an active PSO capacity that enhances the lifespan of SNs. In comparison to the PSO method, the proposed technique improved remaining energy and throughput by 83.79% and 28.90%, respectively. It made use of HSA's superior searching efficiency, which is related to the fact that it develops a new solution from the current one, and PSO's active capacity, which allows it to transfer from one area to another in pursuit of an optimum result, and so it outperformed other algorithms.

V. Sercan [199] compared the SEED algorithm with LEACH and PEGASIS in standings of alive nodes and data packets sent to the BS. In terms of the strength involved, the SEED approach was found to be superior to the other ways, and it also proved to be effective in terms of energy usage. A. Gambhir [37] explained the ABCO-based LEACH algorithm which focuses on the trial in the variability of WSN situations, adjusting the maximum number of rounds (rmax) and the number of SNs (n). For performance estimation, a diversity of factors was considered by the authors. X. Zhao [159] proposed a unique energy-efficient technique called Fitness value-based Improved GWO to expand GWO's search, resulting a more scattering of CHs and a well-balanced cluster structure. SN transmission was updated as per the distance of SNs to the CHs and BS to decrease the usage of the battery. This technique improved the stability period by 31.5% when compared to SEP and 57.8% when compared to LEACH, increasing the data's trustworthiness. When compared to the two algorithms under consideration, the network's throughput was also found to be boosted.

D. Sambo [189] conducted a broad study of proposed optimum clustering systems. They took ten factors into account while evaluating them. The authors provided a comparison of the optimal clustering algorithms based on some characteristics. According to the findings, unified clustering results based on the swarm intelligence standard were found to be more suited for applications that needed low energy ingesting, high data transfer rates, or high scalability. F. Fanian [200] concentrated on this survey for assessing the qualities of various approaches. In terms of clustering characteristics, the compared protocols were divided into macro and micro classes. Authors offer a fresh viewpoint for investigating techniques by taking into consideration practice-based factors, allowing for more rapid comprehension of methodology flaws. R. Sharma [201] concentrated on how various metaheuristic techniques and hybrids perform a critical role in the growth of energy-aware clustering algorithms. The simulation results for remaining energy showed that a hybrid approach to NIAs outperforms standard NIAs techniques in terms of enduring energy per round.

Z. Sun [202] suggested an ant colony optimization (ACO) for WSNs. By incorporating the remaining energy of SNs and the conviction value in which a route path can be formed by comprising two objective functions. To minimize the network's energy ingestion and assure it to regular, the average remaining energy of routing paths has been used by the authors, as the first goal function. To guarantee that the routing nodes are trustworthy, the average trust value was used as another objective function. S. Sankar [203] presented a novel CH selection and cluster building technique based on two stages of the procedure. The SOA's concert is compared to that of IABCOCT, EPSOCT, and HCCHE. According to the simulation results, the advised SOA extends network lifespan by 6-12% and reduces end-to-end latency by 15-22%. P. Nandhini [204] offered a Charged System Search (CSS) and Harmony Search Algorithm (HSA) that examines the difficulties of optimum path selection in WSNs to extend the network's lifetime. Various metaheuristic strategies exist, such as CSS, which can be applied to resolve the routing problem. The approach given in the study work was discovered to be energy efficient and responsive. By improving routing and boosting network longevity, the scheme can choose suitable CHs. The network's total lifetime is increased by the system's selection of efficient CHs with routing optimization.

Above we presented a review of different published bio-inspired algorithms and the contribution of some important algorithms described in Table 4.1 shown below:

Table 4.1. Related nature-inspired algorithms and their contribution

Author	Algorithm	Contribution
P. Kuila [56] **2015**	Particle Swarm Optimization (PSO)	Clustering and energy-efficient routing are two significant WSNs optimization challenges that have been developed using linear and non-linear programming. The lifetime of the network is extended and the energy utilization of the CHs is greatly balanced by taking into account a trade-off between transmission distance and hop count.

T. Bhatia [195] **2016**	GADA-LEACH Hybrid Algorithm	The proposed genetic algorithm-based distance-aware routing system uses GA to optimize CH selection. The suggested method improves CH selection, which extends the network lifetime and rises the amount of data transferred to BS.
Al-Aboody [158] **2016**	Grey Wolf Optimizer (GWO)	The authors suggested a three-level hybrid clustering routing system based on GWO for WSNs. In level one, BS plays a significant role in choosing CHs. In level two, a GWO routing for data transfer is suggested to conserve even more energy. The findings indicated that the suggested method performed better in terms of network lifespan, stability period, and residual energy.
A. Gambhir [37] **2018**	ABC-based LEACH Algorithm	The authors of this study take into account several characteristics, including the number of dead nodes each round, living nodes per round, and packets to BS. Performance metrics outperformed PSO, GA, ACO, and CSA in terms of outcomes.
X. Zhao [159] **2018**	Improved GWO Algorithm	The authors suggested a brand-new WSN method that performs better than the traditional GWO technique. They improved the network's longevity by lowering the average transmission distance and energy usage.
Z. Sun [202] **2019**	Ant Colony Optimization (ACO)	The Pareto multi-objective optimization technique is added to the ACO algorithm in this work to address the WSN routing's resource constraints and security concerns. The simulations performed with NS2 show that the suggested SRPMA can accomplish superior acts in the packet loss rate and average energy consumption.
S. Sankar [203] **2020**	Sailfish Optimization Algorithm (SOA)	To extend the life of the network and reduce node-to-sink latency, this study developed a novel CH selection and cluster building method. According to the simulation results, the suggested SOA extends network lifetime by 5–10% and reduces end–to–end latency by 10–20%.
P. Nandhini [204] **2021**	CS-HAS Hybrid Algorithm	Authors employ CSS to address problems with WSNs' optimum path selection and lengthen network lifespan. The HSA method offers the advantages of easy realization and quick convergence whereas the CSS approach does not require gradient information or continuity in the search space. The study work's method was determined to be effective in terms of energy and network longevity.

4.3 Problem Statement

Every nature-inspired optimization algorithm must strike a balance between global and local search since this is critical for effectively locating the optima. During the early phases of optimization, it is always preferable for the solutions to be encouraged to roam the whole search space rather than congregate the local optima. It is necessary to converge towards the global optimal solution in the latter phases of optimization to identify the optimum solution. Modality of perception in the fundamental butterfly optimization technique, c is a crucial parameter. It will affect the algorithm's performance in two ways.

First, if a big value of c is chosen, it may bypass the most optimum solution early in the optimization process, reducing the algorithm's search performance.

Second, if c is set to a low value, the issue may become trapped in the local optima trap, resulting in premature convergence. The static technique of setting (c) will not be adaptable to complicated real-world circumstances. So, a dynamic and adaptive strategy to change the sensor modality can be used to overcome this issue for global optimization problems.

4.4 Conventional Butterfly Optimization Algorithm

BOA is an advanced meta-heuristic algorithm for wide-ranging optimization that is encouraged by the foraging conduct of butterflies. The coordinated migration of butterflies toward the food source position can be defined as per their behavior. The fragrance in the air is received/sensed and analyzed by butterflies to regulate the possible track of a food source/mating companion as shown in Fig. 4.1. BOA imitates these deeds to locate the hyper search space's optimal location.

(a) (b) (c)

Fig. 4.1. Example of a general social association and behavior of a butterfly with the flowers and environment wherein (a) depicts a butterfly, (b) foraging of the food, and (c) butterfly mating among the flowers [184]

Butterflies are thought to be as confidential as a Lepidoptera which comes under the Linnaean animal kingdom family. Butterflies are found in about 18k different classes all over the world. Their senses are the reason they have subsisted for millions of years. They utilize their five intelligences to discover nutrition and a mate. Navigation from one location to another, hunter invasion, and the decision to lay their eggs at a particular place are also decided with the help of these senses. The fragrance is the important one because it supports butterflies to discover the food, which is usually nectar, even across extensive distances.

Butterflies also employ sense receptors to locate nectar sources. These receptors are distributed throughout the butterfly's body fragments, such as the antennae, limbs, and palps. Chemo-receptors which are sensitive to the chemical in an environment, further help in directing the butterflies to find a suitable fertile and breeding male partner thus ensuring a healthy and strong genetic line. In this process, the pheromone of the female progeny of the butterfly is utilized by the female butterfly to encourage the male butterfly and also helps in identifying them. Each scent in the BOA assembly has its discrete aroma and special touch. To understand how a stimulus is translated into a sensory modality like smell, sound, light, or temperature, it is first necessary to understand how fragrance is calculated in BOA. The complete notion of detecting and distributing the modality is structured around three main terms: sensory modality (c), input intensity (I), and power exponent (a). The natural spectacle of butterflies is formed on two fundamental subjects: the distinction of I and the calculation of F. For leniency, a butterfly's I is connected to a predetermined function. However, f is comparative. In BOA, the fragrance is expressed as a function of the physical intensity of the stimulus as shown in Eq. 4.1.

$$f_i = cI^a \tag{4.1}$$

here (f) is the supposed scale of fragrance, (c) is the sensory modality, (I) is the stimulus intensity, and (a) is the modality-dependent power exponent, which considers for different levels of interest. The global search and the local search are the two most important phases in the algorithm represented using Eq. (4.2) and Eq. (4.3) respectively.

$$x_i^{t+1} = x_i^t + (l\grave{e}vy(\lambda) \times g^* - x_i^t) \times f_i \tag{4.2}$$

here x_i^t is the solution vector x_i for ith butterfly and the current best solution is denoted by g∗. f_i represents the scent of the i^{th} butterfly. Eq. (4.3) can be used to depict the local search phase.

$$x_i^{t+1} = x_i^t + \left(l\grave{e}vy(\lambda) \times x_k^t - x_j^t\right) \times f_i \tag{4.3}$$

here x_j^t and x_k^t are j^{th} and k^{th} butterflies preferred randomly from the solution space. Eq. (4.3) becomes a local arbitrary walk if x_j^t and x_k^t are members of the same sub-swarm. To move from common global search to concentrated local search, IBOA employs a shift probability (p).

Butterfly movement is influenced by the following factors:

- All butterflies are intended to generate a scent that permits them to fascinate one another.
- Every butterfly will soar on an arbitrary track or in the route of the butterfly that emanates the maximum scent.
- The scenery of the goal function affects or regulates the sensory intensity of a butterfly.

4.5 The Proposed Algorithm: Improved version of BOA

IBOA is proposed to lift the network's era and diminish the energy feasting of SNs. Every nature-inspired optimization algorithm must strike a balance between global and local search since this is critical for effectively locating the optima. During the early phases of optimization, it is always preferable for the solutions to be encouraged to roam the whole search space rather than congregate the local optima. It is necessary to converge towards the global optimal solution in the latter phases of optimization to identify the optimum solution.

Modality of perception in the fundamental butterfly optimization technique, c is a crucial parameter. The significance of c can be gauged by the fact that it allows each butterfly in the search process to detect the scents generated by other butterflies and direct the search towards them. This indicates that the more effective the detecting system, the better the findings. The static technique of setting (*c*) will not be adaptable to complicated real-world circumstances. It will affect the algorithm's performance in two ways. First, if a big value of c is chosen, it may bypass the most optimum solution early in the optimization process, reducing the algorithm's search performance. Second, if *c* is set to a low value, the issue may become bounded in the local optima trap, resulting in premature convergence.

4.5.1 Flow Process and Movement of Butterfly

To find the optimum solution, it is essential to converge toward the global optimum solution. Sensory modality © is one of the important parameters (great impact on searching ability) in the basic butterfly optimization algorithm. The better the sensing mechanism, the more efficient the results will be. The static method of setting c will not be adaptive to complex real-world problems. It will affect the performance of the algorithm in two ways.

To increase the effectiveness of the algorithm, the value of c should increase rapidly with a small number of generations while it should increase slowly with a large number of generations. IBOA is modified in such a manner that the butterflies are able to change the value of c dynamically as shown in Fig. 4.2. The flow process for IBOA is as follows:

Step 1: Determine the number of butterflies n, the perceptual constant c, the stimulus factor I, and the switching probability parameter p. Initialize the population using circle chaos mapping.

Step 2: Solve the objective function.

Step 3: Derive the current optimal solution.

Step 4: Randomly select an individual from the current solution.

Step 5: Update the solution.

Step 6: Check the number of iterations. If the number of iterations reaches a certain value, the calculation ends. Check whether the termination condition is satisfied. If it is not satisfied, return to step 3.

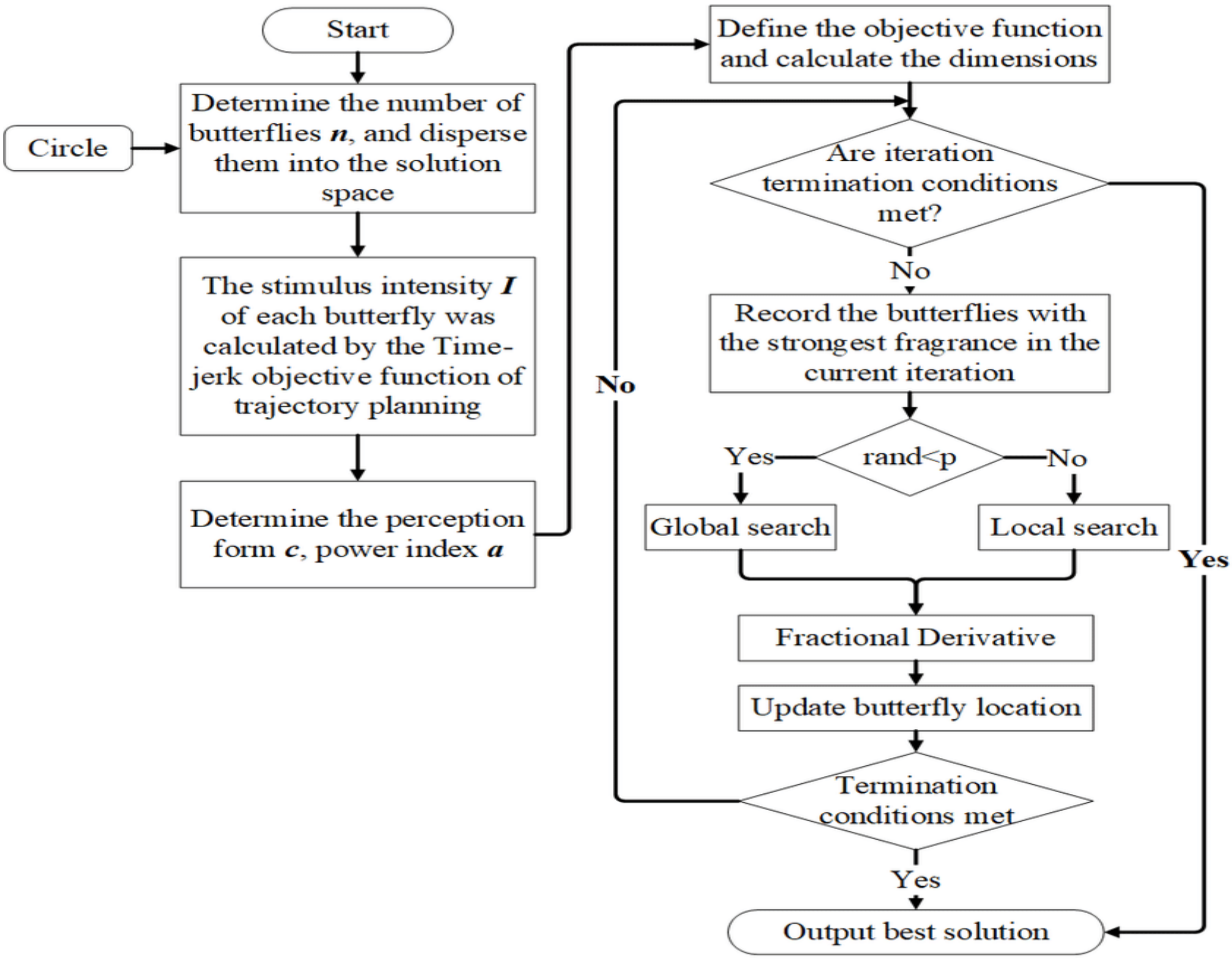


Fig. 4.2. Flowchart of IBOA

4.5.2 Methodology of IBOA

The sensory modality has a significant influence on butterflies' capacity to hunt. With a short number of generations, the value of c should rise quickly, whereas, with a big number of generations, it should increase slowly. The algorithm's efficacy will undoubtedly improve as a result of this. Because of the aforementioned issues and the relevance of the sensory modality parameter, the algorithm has been adjusted such that the butterflies may dynamically modify the value of *c*. So, in this study, a dynamic and adaptive approach to sensory modality is designed and used. The sensory modality (*c*) can be calculated using Eq. 4.4.

$$c^{t+1} = c^t + (0.030/(c^t \times I_{max}) \quad (4.4)$$

here t is the current no. of iterations and I_{max} is the maximum no. of iterations. Instead of levy flights, pseudorandom numbers are employed in this investigation. The global and local search phases of the proposed IBOA are defined in Eqs. (4.5) and (4.6), respectively, to take into consideration the foregoing arguments.

$$x_i^{t+1} = x_i^t + (r^2 \times g^* - x_i^t) \times f_i \quad (4.5)$$

here x_i^t is the solution vector x_i for ith butterfly and the existing finest solution discovered among all is denoted by g∗. f_i represents the scent of the i^{th} butterfly while r is a random number in [0, 1].

$$x_i^{t+1} = x_i^t + \left(r^2 \times x_k^t - x_j^t\right) \times f_i \quad (4.6)$$

here $\boldsymbol{x_j^t}$ and $\boldsymbol{x_k^t}$ are j^{th} and k^{th} butterflies. Equation (4.6) becomes a local random walk if $\boldsymbol{x_j^t}$ and $\boldsymbol{x_k^t}$ are members of the same sub-swarm and r is a random choice in the range [0, 1].

4.6 Simulation Results and Comparative Analysis

Extensive simulations were run to assess the performance of IBOA. MATLAB was used to create both the network model and the IBOA technique. Under the same conditions, the IBOA algorithm's results are compared with ZSEP and LEACH. Table 4.2 lists the most important simulation parameters of the performed simulations.

Table 4.2. Simulation parameters utilized in this study

Parameter	**Value**
Area	500x500 m^2
Number of sensors	100
Number of clusters	10
Base station location	250,250
Protocol	Low energy clustering protocols
Mobility	Random
Initial Energy	0.5 J
Transmission Energy	50 x 10^{-9} J
No. of iterations	500

Table 4.3 shows the concrete values of throughput, first node dies (FND), and elapsed time. From the results, we can conclude that IBOA throughput is increased by 480.09% and 26.53% in comparison to LEACH [192] and ZSEP [193] respectively.

Table 4.3. Comparison of network throughput, first node dies, and elapsed time

Algorithms	Throughput	First Node Die (FND)	Elapsed Time
LEACH [192]	0.3503	655	0.000564
ZSEP [193]	1.6059	986	0.000516
IBOA	2.0320	1345	0.000414

Fig. 4.3 shows a distributed wireless sensor network including BS and multiple SNs. The simulation results of the anticipated network model for sensor deployment using IBOA depicted in Fig. 4.4. The no. of alive nodes is more with the application of the proposed IBOA for a given no. of rounds. In case of LEACH, all nodes are alive up to 655 rounds with an elapsed time of 0.000564, and ZSEP takes 986 rounds for FND with an elapsed time of 0.000516. IBOA significantly increased the throughput, takes 1345 rounds for FND with an elapsed time of 0.000414. Fig. 4.4. also shows comparison of performance of the IBOA to different optimization techniques.

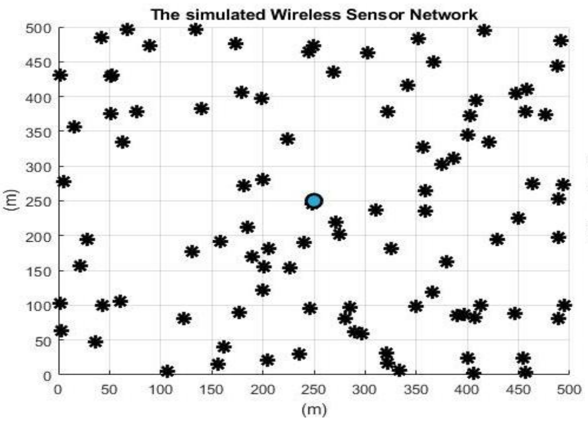


Fig. 4.3. Distributed wireless sensor network including BS and multiple SNs

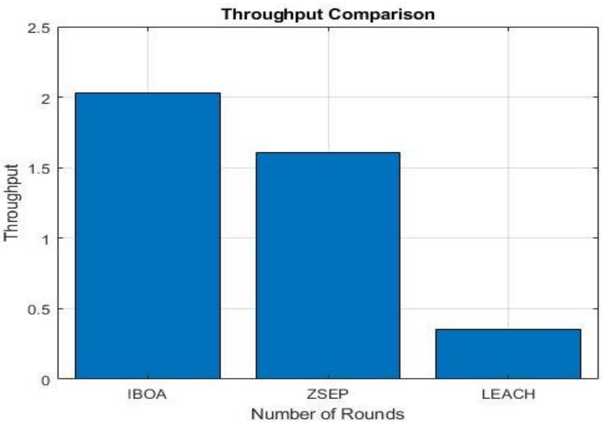
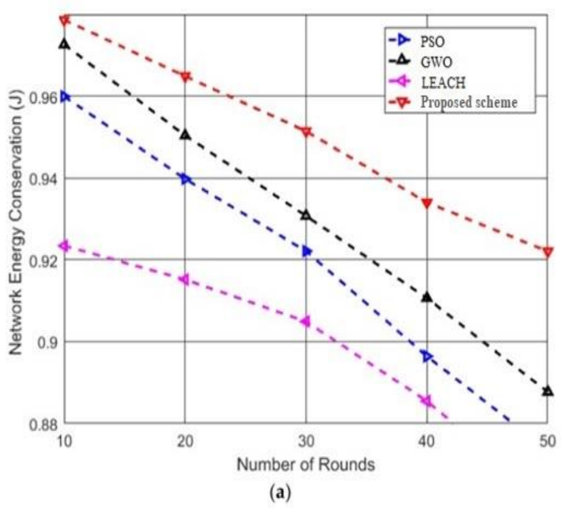


Fig. 4.4. Comparison of performance of the IBOA to different optimization techniques

The no. of dead nodes is higher in LEACH [192] protocol as compared to PSO [75], GWO [28] and IBOA. This shows that the network's lifetime as shown in Fig. 4.5 is better than with the proposed IBOA.



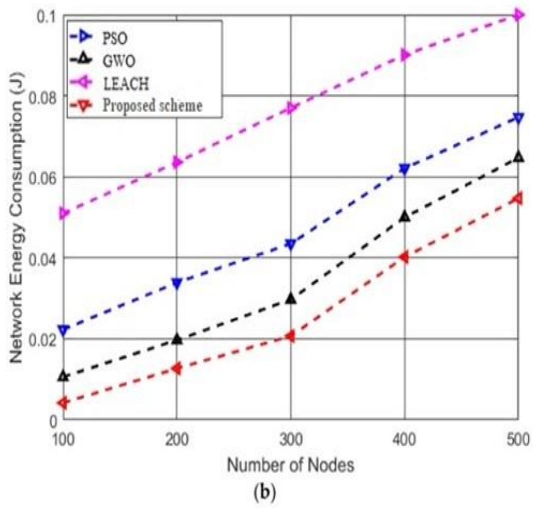
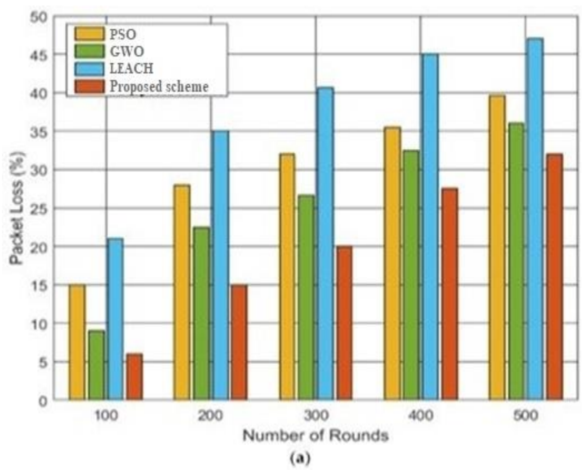


Fig. 4.5. Network energy conservation with respect to the number of rounds

Fig. 4.6. Network energy consumption corresponding to the number of nodes

The packet delivery ratio is improved using the new approach IBOA as shown in Fig. 4.6 for the same no. of rounds and CHs. This shows that IBOA is more effective for delivering more no. of packets to the BS. The throughput and convergence rate of improved BOA is also much better than PSO [75], GWO [28] and LEACH [192] as shown in Fig. 4.7 and Fig. 4.8.



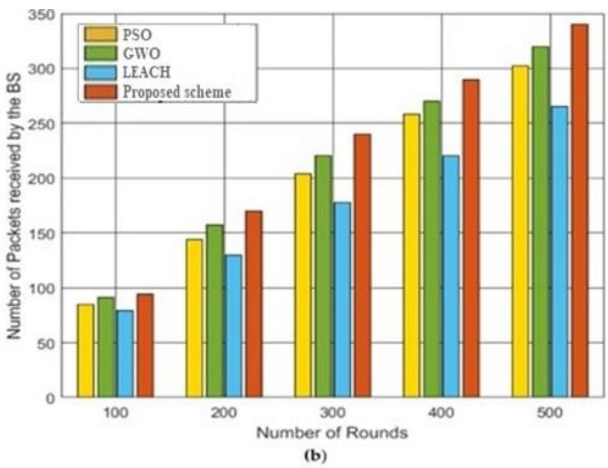


Fig. 4.7. Packet loss with respect to the number of rounds

Fig. 4.8. Number of packets sent to BS with respect to the number of rounds

The anticipated algorithm attained a better throughput, network lifetime, and packet delivery compared to [28], [75], [192], and [193]. Concerning the proposed and existing algorithm performance, the following metrics are purposed for calculation.

- Residual energy: This comprises each node's average remaining energy as well as the energy differential between the most and least energetic nodes.
- No. of alive nodes: This includes remaining alive nodes after a fixed number of iterations.
- Packet Loss: Difference between of amounts data bundles that BS has received and transmitted by SNs.

4.7 Conclusion

For global optimization problems, a variable sensor modality improved butterfly optimization (IBOA) technique is given in this section. The suggested technique employs a dynamic and adaptive strategy to transform the sensor modality, which was set to a constant value in the conventional butterfly optimization algorithm. The butterflies' seeking skills were improved by the varying value of sensor modality. The results showed that the butterflies utilize their knowledge more effectively in the suggested algorithm to execute exploration and exploitation more efficiently than in basic BOA. The comparison is done based on given parameters with existing protocols like LEACH, PSO and GWO. Result depicts that in the 300th round, the packet loss for LEACH is 42%, which is double the loss for the proposed scheme, i.e., 21%. Moreover, GWO and PSO show 5% and 12% more packet loss when compared with the proposed scheme. Only unconstrained issues are studied in this work; however, it will be fascinating to observe how well the modified butterfly optimization approach performs on restricted problems and remains the future scope of this work.

Chapter Related Publication

- **Rajiv Yadav,** Indu Sreedevi, Daya Gupta, "Performance Analysis of Energy-Efficient Cluster-based Routing Protocols with an Improved Bio-inspired Algorithm in WSNs," International Conference on Data Analytics and Computing, **Springer, (Scopus Indexed),** China, 2022.

Chapter 5

Energy-efficient Feature Selection and Classification Techniques in WSNs for IoT Applications

5.1 Introduction

WSN is an assembly of robust SNs that exchanges data through a wireless system. A sensor is a device that collects information about a physical thing or an incident. The sensor center is made up of recognition, actuation, and power elements that are all synchronized on one or more sheets that make up the embedded system. A WSN employs several sensors to communicate data to a designated location over the internet [38]. The SN's sensor element detects the observable status of a scenario or physical environment, while the microprocessor guarantees that the data is effectively calculated. The node's wireless radio also takes guarantees communication with nearby nodes. WSNs are frequently installed in distant, rugged, and unsupervised locations for extended periods. Node preservation after the deployment of sensors is difficult due to the inaccessibility of these locations. Temperature, sound, vibration, and weight are all detectable by them. The application of WSNs is becoming widespread with the growing use of the IoT, supervision, and distant monitoring of mobile phones [205], [206].

WSNs are frequently employed for a variety of purposes, such as covering military applications, environmental catastrophes, animal monitoring, safety inspections, etc. [207]. The basic goal of WSNs is to assemble useful statistics in a steady situation. Because there are multiple barriers to the security of these networks and IoT, providing and sustaining data protection is a critical concern. Distributed Denial of Service (DDoS) assaults, server negotiation caricature attacks, and protocol-specific attacks are all examples of these types of attacks. A DDoS assault can swiftly deplete available resources through a variety of methods, including flooding [208]. The risk is far higher when it comes to military and medical uses. Presenting a security model that considers these limits while providing security is a key task nowadays. Extortion has been recognized as one of the key reasons for such assaults, which have continued to grow in magnitude and sophistication.

Various Intrusion Detection Systems (IDSs) have been suggested by scholars to safeguard WSNs from adversary intrusion [209]. The first uses signatures of known attack patterns, whereas the latter uses data mining, Machine Learning (ML), and statistical modeling to profile a statistical use model. Signature-based detection has a noteworthy weakness that cannot detect unknown assaults. As a result, a hybrid approach has emerged, which combines the complementing features of both methodologies to attain a greater detection rate.

Feature Selection (FS) helps to improve simplification by decreasing over-fitting, computational efficiency, and stuffing, as well as improving cataloging precision by developing an accurate prediction model as shown in Fig. 5.1.

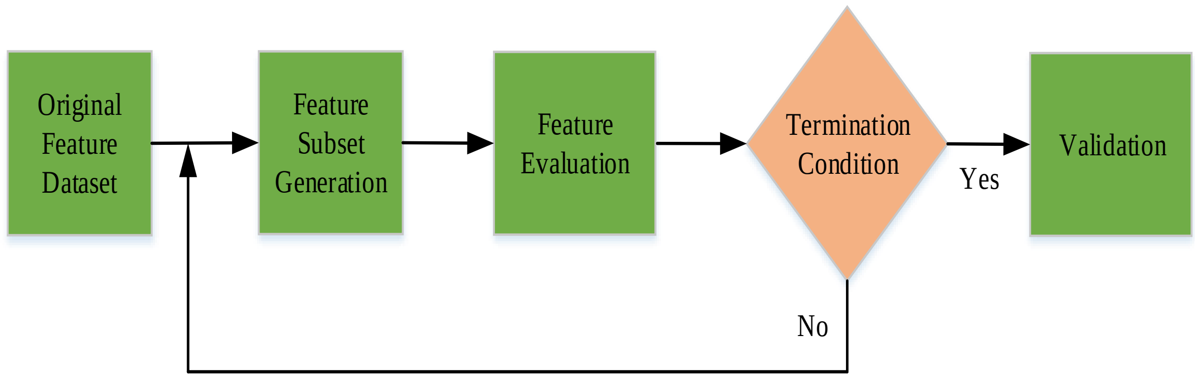


Fig. 5.1. Feature selection process

The rapid increase in advanced technologies, like mobile communications, device availability, and processing systems, has led to the next era of interaction known as the IoT. As a result, IoT security is a serious concern to safeguard the hardware and networks of IoT systems [8]. Security has not been considered in the creation of networking gadgets because they are still relatively recent inventions. Academic research on confidentiality problems for IoT devices has shown encouraging breakthroughs in recent years. The tactics and security processes now available are primarily based on classic network security approaches. Establishing security mechanisms in an IoT system is more complex than in a traditional network due to the wide variety of technologies and devices, as well as the scalability or number of nodes in the network.

WSN has vast real-world applications for IoT [210] in the following areas:

- Defense: In the battlefield for surveillance and monitoring.
- Environment: Forest fire detection, Pollution monitoring, and early detection of earthquakes.
- Health Care: For patient diagnosis and monitoring.
- Agriculture: Crop monitoring, soil monitoring, temperature, and humidity.
- IoT: Home automation, Smart cities, and smart grids.

In this chapter, we introduce an FS phase using a Fast Correlation-Based Feature Selection (FCBFS) method to select almost one-fourth of the original dataset (10 most important features) before classifying it with different classifier algorithms. Because SNs have a limited number of resources, feature selection is used in ML approaches to minimize duplicate characteristics and lower the difficulty of the projected scheme. The IDS benchmark dataset, NSL-KDD, with 41 characteristics, was used to test the performance taking into account precision, classification accuracy, recall, and F-score [211]. In comparison to [212], the findings reveal that our suggested strategy successfully reduces the number of characteristics while retaining excellent classification accuracy and precision, as well as F-score.

5.2 Related Work

A deep learning strategy built on the DNN technique was put into practise by Potluri and Diedrich [96]. They exclude 41 characteristics from consideration, of which just 27 were used in the experiment. Mixed outcomes have been obtained. In addition to the RNN method, You et al. [213] applied a deep learning strategy. On the well-known dataset, experiments were conducted. SVM-based and Naive-based methodologies underwent comparative study. With a 92.7% accuracy rate, they reported to have attained encouraging findings. Alrawashdeh and Purdy [100] put unsupervised feature reduction-based deep learning to use. For the experiment, the logistic regression classifier was utilized. This study employed the well-known dataset KDD Cup 99, and the findings showed a 97.90% detection rate. Ding and Wang [214] combined the traffic anomaly approach with deep learning techniques. They made an effort to fix the dataset's issues. They also stated that they had had encouraging outcomes.

A highly effective approach to identify any network misbehaviour was presented by Attia et al. [215]. They showed how the IDS system's planned mechanism differed from the system's current mechanism. Their model's detection rate is 95%, while its false detection rate is around 5%. Gunduz and Ceter in [216] compared the outcomes of four different categorization methods. In this study, many algorithms like fuzzy unordered rule induction algorithm (FURIA), decision trees, multilayer perceptron network, and support vector machines were applied. They came to the conclusion that the best-first search method (BFS) and CFS provide the model with the best characteristics. They used classification techniques to choose the top 11 features out of the KDDCUP99 dataset's 41 features. FURIA produces the best outcomes.

It was recommended by Aburomman and Reaz in [217] that not all characteristics from the dataset had to be included to create the model. They evaluated the attributes of their model utilizing gain ratio, Chi-square-based attribute selection, symmetrical uncertainty feature selection, and Naive Bayes classifier before eventually classifying the case. Using classifiers like Naive Bayes, Random Tree, Meta Tagging, Decision Stump, Ada-BoostM1, Pegging, J48, and REP Tree, Pande et al. in [218] suggested an approach utilizing hybrid algorithms. To filter the data on the NSL-KDD dataset, they employed the Vote algorithm with information gain. Using a binary class, they achieved 99.81% accuracy, and with a multiclass dataset, 98.56% accuracy.

Deep learning techniques were applied for the health monitoring device by Zhao et al. in [219]. We examined and evaluated four deep learning techniques. On a well-known dataset, the experiment was run, and it produced promising findings. The dataset's top features were chosen using the GWOSVM-IDS approach by Safaldin et al. in [95]who then created a support vector machine classification model using the modified binary grey wolf optimizer. To ensure the transfer between packets and sensors was safe, they used a high-level security mechanism. This model has performed well when categorizing assaults.

5.3 Problem Statement

Numerous studies have demonstrated that the accuracy and general performance of the model can be improved by using fewer redundant datasets or by choosing pertinent characteristics from a dataset. Scholars have used a variety of approaches for selecting features to accomplish this, and these investigators have also combined different techniques as well as feature rating methods to identify the dataset's most crucial characteristics. Many attacks like DoS, Probe, R2L, and U2R are the main challenges in the security of WSNs. Several investigations have demonstrated that removing redundant information from a dataset can improve the precision and effectiveness of a model. Therefore, we propose an FCBFS method with XG-Boost for NSL-KDD intrusion detection benchmark dataset to address the threats and security issues in a complicated WSN for IoT applications. Evaluation metrics such as accuracy, precision, recall, and F1-Score have been calculated to gauge the performance and robustness of the proposed research work.

5.4 Methodology

5.4.1 Dataset Description and Pre-processing

The collection of significant features from the dataset is critical to the performance of any ML model. The most common NSL-KDD dataset is utilized for classification in the implementation. There are 41 distinct attributes in this dataset, including content type, basic kind, and traffic type as shown in Table 5.1. Each record reveals diverse flow aspects, each of which is categorized as an attack type or normal.

Table 5.1. Dataset description [118]

Criteria	**Description**
Dataset	NSL-KDD
Number of records	1,49,470
Number of attack categories	4 (DoS, Probe, R2L, U2R)
Number of network features	41
Type of network features	Basic features, traffic features, Host features and content features.
Number of records in the training set	1,26,620
Number of records in the test set	22, 850

The 42nd attribute offers evidence of the five unlike types of system connection vectors. They are further separated into one regular class and four attack classes. DoS, Probe, R2L, and U2R are the four attack categories which are divided as the four assault types [220]. The attack classes are described below:

- **Denial-of-Service attack (DoS):** The earliest type of cyber squeezing assault is a DoS attack in which the invader makes the server extremely busy, and as a result of that program, the legitimate user is denied access. This type of assault often employs the flood attack approach, which involves making unwarranted system and network requests, causing the computer system or network service to refuse the targeted person's fair service. In a DoS attack, the victim's resources are depleted, rendering it unable to handle genuine requests - for example, syn flooding.

- **Probe:** This attack is an effort to fold data about a system of PCs with the clear goal of evading its security protections. When the attacker links to the associated targeted device, the scanning traffic supplied by the attacker can scan the computer for vulnerabilities, making it easier to choose acceptable means and objects for further assaults. Surveillance and other probing attacks' objectives are to gain data about the remote target e.g., port scanning. For instance, SATAN, SAINT, port weep, and so forth.

- **Remote to local attack (R2L):** This attack entails sending particular information to the target across the network, then obtaining approval to access the victim's device to gain access to the victim's terminal structure for further assault. The directory contains a directory with an extremely lengthy name; the FTP client then executes (unintentionally) one or more commands included in the name with the user's permission. The attacker encroaches into a distant appliance and acquires local access to the victim mechanism through an illegal entree from a remote appliance. For example, password guessing.

- **Unauthorized access to local superuser (root) privileges (U2R):** A regular account is utilized by a U2R attacker to log into the target system. The attacker attempts to obtain root/administrator rights by manipulating the softness of the victim's regular account. This strategy is similar to going backward. After the attacker has access to the victim's terminal structure, he will utilize the attack circulation to modify his rights from regular users to system access rights and then eventually gain administrator or root credentials. It gains access to the system's roots and launches a series of assaults and illegal efforts such as various Buffer overflow attacks, Perl, rootkit, etc. Relevant features: "number of file creations" and "number of shells prompt invoked,".

The most important phase in the experiment is the model step. Because the goal of this study is to treasure the best FS technique and ML model for dealing with the challenges, as shown in Fig. 5.2.

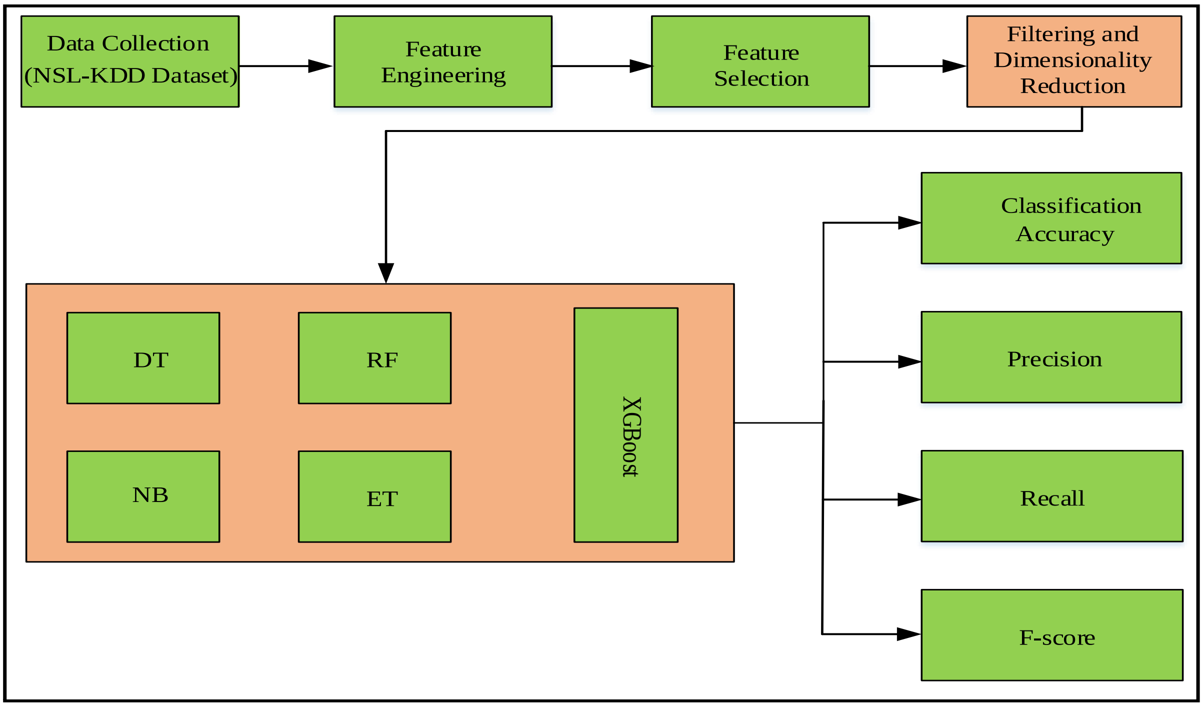


Fig. 5.2. Experimental flow of proposed work

Data pre-processing is the most significant phase in every ML project. The basic goal of data cleaning is to assure data accuracy by searching for duplicated or unrealistic information. However, there are numerous non-numeric properties of data collection. In general, Scikit-Learn classifiers can handle numeric characteristics well but not non-numeric attributes. Rather than merely converting non-numeric qualities to numeric attributes, it is important to use a suitably specialized encoding mechanism to preprocess the data in the data processing.

The computer cannot match one attribute to another since they are all discontinuous and unordered, but it may place N potential outcomes of a conceivable attribute in N registers using binary to store the data after applying a certain encoding method. Each state has its own set of register bits, with only one active at any one moment. In this experiment, one-hot coding is used. In preprocessing, the raw dataset is cleaned to convert it into a set-up that can be utilized to train an ML model. When creating a well-organized model, the data must be in a precise format.

5.4.2 Feature Selection

The goal of FS is to eliminate material that is redundant or useless. Not every data aspect makes sense. This is a method for choosing a subset of important attributes that accurately describes a situation. In addition, the selection will reduce the amount of data picked. It's critical to choose the right characteristics from the complete dataset, as shown in Table 5.2, because they can impact the model's performance and accuracy.

Table 5.2. NSL-KDD dataset features

No.	Dataset features	No.	Dataset features	No.	Dataset features
1.	Duration	15.	Su_attempted	29.	Same_srv_rate
2.	Protocol_type	16.	Num_root	30.	Diff_srv_rate
3.	Service	17.	Num_file_creations	31.	Srv_diff_host_rate
4.	Flag	18.	Num_shells	32.	Dst_host_count
5.	Src_bytes	19.	Num_access_files	33.	Dst_host_srv_count
6.	Dst_bytes	20.	Num_outbound_cmds	34.	Dst_host_same_srv_rate
7.	Land	21.	Is_host_login	35.	Dst_host_diff_srv_rate
8.	Wrong_fragment	22.	Is_guest_login	36.	Dst_host_same_src_port_rate
9.	Urgent	23.	Count	37.	Dst_host_srv_diff_host_rate
10	Hot	24.	Srv_count	38.	Dst_host_serror_rate
11.	Num_failed_logins	25.	Serror_rate	39.	Dst_host_srv_serror_rate
12.	Logged_in	26.	Srv_serror_rate	40.	Dst_host_rerror_rate
13.	Num_compromised	27.	Rerror_rate	41.	Dst_host_srv_rerror_rate
14.	Root_shell	28.	Srv_rerror_rate		

To begin with, unconnected traits may signal some level of effectiveness to a certain situation, but this does not imply that they are inextricably linked. As a result, suitable models cannot be built appropriately. This will not only damage the data's correctness and model's dependability, but it will also cause overfitting issues. Second, the bigger the measurement of the source information, the lengthier the dispensation and computing time for the classifier will be. The innovative data includes a large number of dimensions, which increases ML time, lowers efficiency, and diminishes prediction efficiency. The impact of time on intrusion detection difficulties is unbearable.

Reducing duplication and picking more appropriate characteristics aids in:

- ❑ Reduces the dataset's size.
- ❑ Reduces the likelihood of data being manipulated.
- ❑ Reduces the chance of overfitting.

- Reduces the amount of time it takes to train the model.
- Enhances the model's accuracy and reliability.

The most crucial stage is to remove unnecessary, redundant characteristics from the dataset. Studies suggest that the development of ML models using irrelevant characteristics or feature affect the evaluation metrics of the model. In this paper, we will use FCBFS method that automatically chooses the 10 most relevant features from the 41 features of the NSL-KDD dataset. Information Gain (IG) is one of the filter FS strategies used in selecting appropriate qualities from a set of features. When the value of a feature is unknown, IG reduces the uncertainty associated with determining the class attribute. It is based on information theory, and it is used to rank and choose top features to minimize feature size before the learning process begins. The information gain value of the dataset's features is rated. A characteristic with a low information gain value indicates that it has little impact on data categorization. As a result, features with small information gain values can be discarded without having a significant impact on the classification model's performance. Eq. (5.1) may be used to compute the information gain value for class C and input feature X.

$$IG(C,X) = H(C) - H\left(\frac{C}{X}\right) \tag{5.1}$$

Here, H(C|X) is the probability of class C for a given feature X, while H(C) is the entropy of class C.

The dataset is first subjected to the FCBFS approach. Pearson's correlation coefficient (PCC) is used to create a correlation matrix. The PCC is derived by dividing the covariance of two values and then multiplying the divided value by the standard deviation of each value, as shown in Eq. (5.2).

$$Coefficient = \frac{covariance(x,y)}{(stdv(x) * stdv(y))} \tag{5.2}$$

Here, x represents data and y is a random variable.

The link between the attributes may be understood using these coefficients. The coefficients' values are always between -1 and 1. A symmetric matrix is formed by evaluating this connection and placing those standards in that matrix. Correlating characteristics are discarded as a consequence of this observation. Before eliminating the linked features, there were 41 features; after lowering them, only ten features remained.

5.4.3 Feature Scaling

As the feature might comprise big values, the model must be trained on larger standards, which can have a greater impact on the outcome; feature scaling is a necessary step in creating an effective ML model. Pre-processing just prepared the data for input into the model, whether the model's conclusions are valid. Whether the subsequent states can be predicted properly is dependent on the data scaled [220]. Each piece of data contains numerous properties, yet when transformed into a number, each feature has a different assortment of standards. The normalizing technique is used to scale features. It is a method of transforming data points into the 0–1 range while maintaining their relationship. Data with a greater worth has a sophisticated percentage, which has a direct impact on data training accuracy. As a result, you must average each of these values, then subtract the mean from the eigenvalue before dividing by their standard deviation. This phase is critical for improved accuracy, and it must be completed before the model is trained.

5.4.4 Classification Techniques and Execution Process

The classification phase is the last step in the development of the intrusion system. In this step, a labeled training dataset is used for experimentation and training of the classifier with known labels. Later the effectiveness and robustness of the classifier are checked upon the testing dataset, which doesn’t contain obvious labels. The training and testing stages of the model that employ classification-based algorithms can be separated into two categories. Labeled data are used in the training step to acquire a specific classifier. This classifier may then be used in the test step to determine if a test instance is normal or anomalous. Some classifiers like DT, RF, NB, ET, and XG-Boost are used in our experiment with the NSL-KDD dataset [221] [124].

5.4.4.1 Decision Tree (DT) Classifier

To forecast the value of the target variable, a DT classifier constructs a model that acquires basic decision rules using data attributes. A source node and many leaf nodes make up a DT. Leaf node relates to decision outcomes, whereas other nodes agree to feature tests. The sample set is partitioned into child nodes based on the outcome of the attribute test. The whole set of examples may be found at the root node. Each node's route from its root to its leaf resembles a series of choice tests. The goal of DT learning is to create a high generalization ability that can process unknown outcome samples. The DT model should be used in combination with entropy to be effective. The training set's entropy value (H) is computed in bits, as shown in Eq. (5.3) and may be written as:

$$H = -\sum_{i=1}^{n} P(C_i)\, log_2\, P(C_i) \tag{5.3}$$

Here, H is the entropy and P is the probability for different samples.

The DT classifier does not adjust to tiny data changes. A little change in the data provided to the classifier might cause the DT structure to become exceedingly unstable. They have a lesser level of accuracy when dealing with similar types of data. If the data is connected or ambiguous, determining the nodes is extremely difficult. It is not appropriate for issues in which there is very little information about the data. DT, like IDS, cannot be used to detect assaults that aren't known ahead of time.

5.4.4.2 Random Forest (RF) Classifier

RF is a classification approach based on ensemble learning. The method employs the self-sampling strategy to build a fresh training sample set for DT by randomly extracting N samples from the old training set with a retrieval repeat. Then repeat this stage to build M more DTs at random; this is how the RF method got its name. The number of votes in the classification tree is used to determine the results of multiple decision trees. RF classification has several advantages, including improved classification accuracy, a novel technique for discovering significant variables, and the capacity to function in a complicated interaction model, including statistical data analysis and a missing value algorithm. Even huge datasets with many characteristics may be exploited using RF. It determines the value of each characteristic. It can handle imbalanced datasets and does not overfit the data. It is difficult to comprehend. Real-time categorization prediction becomes tougher as the number of trees grows. When a huge number of trees are created, the model becomes slower.

5.4.4.3 Naïve Bays (NB) Classifier

The Bayes theorem states that the probability of occurrence of one event is distinct from the other. The NB classifier is a probabilistic model based on the Bayes theorem. The approach is based on the assumption that the existence of any feature in the class is independent of the presence of any other characteristic. The possibilities of all class characteristics lead to categorizing unknown data into specified classes on their own. The Bayes theorem is used to calculate the prediction error of the class based on the probability of the predictor variable P(c/x). Class C is unaffected by the probability of any of the other variables in the class. The NB method is a scalable and fast classification technique. It may be used to solve problems with binary and multi-class categorization. It's simple to set up since it relies on basic counts of class conditional independence. With a tiny dataset, it works great. The NB technique cannot draw any link between the properties of the class since it is based on the premise that they are independent of one another. The algorithm implementation becomes more difficult when dealing with massive datasets.

5.4.4.4 Extra Tree (ET) Classifier

Finally, there is an additional tree classification model that is similar to the RF model but varies in two aspects. First, the additional tree uses all training samples to create each DT, i.e., each DT uses the same training data. Second, to realize the DT's bifurcation, the RF selects the optimal junction attribute, and the additional tree selects the bifurcation value in a fully random manner. When a numerical value is used for the feature property, an arbitrary number between the extreme and least value is chosen at random. The left branch is considered when the sample's feature attribute value is larger than the value, and the right branch is considered when it is less than the value. As a result, the sample is divided into two branches based on the characteristic property. The divergence values of all the feature traits in the node were acquired using the aforesaid approach after negotiating all of the feature attributes in the node. To achieve the node's bifurcation, we choose the form with the highest bifurcation value. This procedure is more arbitrary, implying that it has a greater ability to explore.

5.4.4.5 XG-Boost Classifier

XG-Boost is a popular regularizing gradient boosting framework is offered by the open-source software package for classifying huge datasets in a short period. There are several compensations to this algorithm, which contribute to its current admiration. It uses parallel processing, which allows customers to acquire their findings faster. We can compare performance and find that the XG-Boost algorithm is superior to other algorithms in terms of accuracy and false positive rate. Additionally, XG-Boost can detect changes in the high-speed internet environment very fast and make adjustments. The most crucial aspect of XG-Boost is its capacity to scale with its growing network size. The classifier's parameters can also be tweaked to improve the results. XG-Boost's target function is specified in Eq. (5.4) given below:

$$F(\phi) = L(\phi) + \Omega(\phi) \tag{5.4}$$

$$\text{where } L(\phi) = l(\hat{y}_i \, y_i) \tag{5.5}$$

$$\Omega(\phi) = \Upsilon T + \frac{1}{2}\lambda\|\omega\|^2 \tag{5.6}$$

Here, ϕ is a term that refers to the formula's different parameters, $L(\phi)$ is a differentiable convex loss function that trials the variance between the prediction $\hat{y}_i$ and the target y_i as shown in Eq. (5.5). $\Omega(\phi)$ is a standardized term that fines complex models as shown in Eq. (5.6) in which T is the number of leaves in the tree, and Υ is the learning rate which ranges from 0 to 1, ω is the weight of the leaves, and λ is a regularized parameter. Increase this item's value to further inhibit fitting and improve the model's generalization capacity. The order in which all of the FS strategies are used is crucial since the accuracy of the models is affected by picking characteristics in various combinations. Furthermore, these FS strategies are organized in such a way that the model's calculation time may be reduced. The calculation time will gradually rise if the sequence of FS techniques is changed.

5.4.5 Prediction and Evaluation

This step utilizes a test dataset to make estimates for the model that has been created, and then cross-validation is performed. To show the findings and comparison of the classifiers, performance measures such as accuracy, precision, recall, and f-score are evaluated. Each of these evaluation parameters is obtained from the four basic features of the confusion matrix, which reflect the real and predicted groups. We can easily assess the prediction performance of each model for different types of attack activity in the training set and test set after training by looking at the matrices below. The performance of proposed methodology is measured by using a confusion matrix with various classes of attacks including DoS, Probe, R2L, U2R. The predicted model gives better results by training the NSL-KDD dataset using XG-Boost classifier.

The following metrics will also be used to evaluate the efficiency of our proposed solution.

- **Classification Accuracy:** The percentage of properly defined data from the complete set indicated by the TP and TN condition or the percentage ratio of correctly specified data to the full dataset is known as classification accuracy, as shown in Eq. (5.7).

$$\text{Accuracy} = \frac{T_p + T_n}{T_p + F_p + F_n + T_n} \tag{5.7}$$

  Here, TP is number of true positives, TN is number of true negatives, FP is number of false positives, and FN is number of false negatives.

- **Precision:** The uniformity of the measuring findings is referred as precision shown in Eq. (5.8).

$$\text{Precision} = \frac{T_p}{T_p + F_p} \tag{5.8}$$

- **Recall:** The proportion of similar occurrences recovered is known as recall, shown in Eq. (5.9).

$$\text{Recall} = \frac{T_p}{T_p + F_n} \tag{5.9}$$

- **F-score:** The weighted harmonic mean of precision and recall is the F-score shown in Eq. (5.10).

$$\text{F} - \text{score} = \frac{Precision \times Recall}{Precision + Recall} \tag{5.10}$$

5.5 Experimental Results

5.5.1 Quantitative Analysis

First, using various assessment and prediction factors, we arrived at the results where the performance of each model in the training set and test set can be predicted clearly. The inferences achieved from the proposed FS process helped in determining the ranking of the most relevant features. They were further chosen and acted as input to the ML models, which classified them into normal or anomaly classes of traffic packets. Using our proposed approach, the FS procedure is conducted to identify the ten highest-ranking features that make up one-fourth of the NSL-KDD dataset. In this study, the training and test sets, KDDTrain.csv and KDDTest.csv, respectively, had 126,620 and 22,850 instances. Every classifier included in the model is trained many times with various combinations of features. Additionally, 8% of the training data and 20% of the testing data are used to validate the complete assessed model. The results produced by this approach will always be less biased, making it the most trustworthy validation technique.

The experiment is run on a work station (GPU) with 128 GB (127 GB usable) of RAM and Intel(R) Xeon(R) Silver 4214 CPU @ 2.20 GHz 2.19 GHz (2 processors). Following a quick correlation-based FS on the dataset, the model is trained using each classifier utilizing the top 10 automatically determined features. For our experimental study, we employ Python software[222], and an ML tool composed of several ML algorithms. Python's arguments are set to their default settings during categorization. Utilizing an open-source NSL-KDD dataset, we evaluate the effectiveness of our suggested technique. The fact that NSL-KDD is open source and easily accessible online is what drives its use in our work. The outcomes of using the FCBFS approach combined with various training models are shown in Table 5.3.

Table 5.3. Performance metrics with different classifiers

Classifiers	Accuracy	Precision	Recall	F1-Score
Decision Tree	99.74	99.72	99.74	99.72
Random forest	99.75	99.72	99.75	99.73
Naïve Bayes	97.52	52.46	95.70	69.64
Extra Tree	99.67	99.64	99.67	99.66
XG-Boost	99.84	99.83	99.84	99.82

The classification accuracy, precision, recall, and F-score of our technique with ten features have improved. Fig. 5.3 displays the performance metrics for all five classifiers using the FS approach.

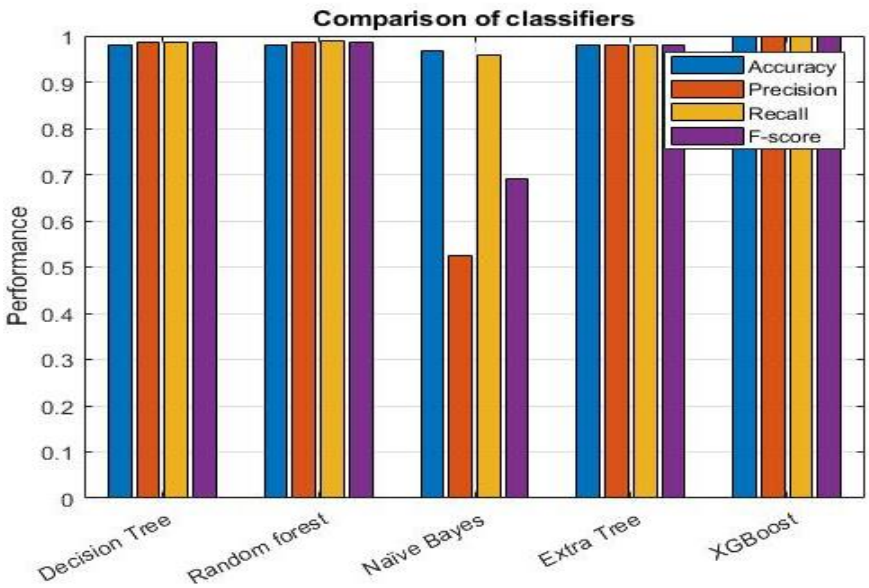


Fig. 5.3. Performance metrics comparison

To produce a significant accurate and minimal spurious item through the regular repetition of the model, the XG-Boost classifier combines a large number of tree versions with lower distinguishing reliability. Although XG-Boost uses fewer resources than current approaches, it can scale up to billions of successful samples. Furthermore, it can be approximated on the out-of-core, conserving memory on the controller.

Results from five different training methods were compared, wherein it was found that FCBFS with the XG-Boost classifier was producing higher positive outcomes across the board. The classification accuracy for our proposed method with different classification techniques is also shown in Fig. 5.3. The qualitative analysis in the next section includes the accuracy comparison of different existing classifiers with our proposed method's results.

5.5.2 Qualitative Analysis

The model's accuracy, precision, recall, and F-score values are computed for all five classifiers using the FCBFS approach. Comparing our suggested strategy to the entire feature set and individual filter selection methods, we find that it takes the least amount of time to create the model. Table 5.4 provides a comparison of the acquired accuracy with other methods currently in use. As a result, our FS approach is effective and simple.

Table 5.4. Comparative analysis of proposed technique with existing methods

Year	Author's Name	Feature Selection	Classifier	Accuracy (%)
2015	D. Deshmukh[223]	FCBFS	NB	88.20
			NB Tree	94.60
2015	B. Ingre[224]	CBFS	ANN	97.50
2016	K. Kumar[118]	CBFS	NB	97.52
		IG	NB	97.11
2016	O. Osanaiye[117]	Ensembled	DT	99.67
2017	B. Subba[116]	PCA	SVM	99.13
			DT	96.85
2018	H. Benaddi[225]	PCA	k-NN	94.00
2019	M. Amir[108]	K-mean with PCA	KNN	95.50
			RF	96.66
2021	A.Thakkar[124]	IG, Chi-square, RFE	SVM	99.60
2022	**Proposed Technique**	**Fast Correlation-based feature selection (FCBFS)**	DT	99.74
			RF	99.75
			NB	97.52
			ET	99.67
			XG-Boost	**99.84**

As presented in Table 5.4, the proposed technique with the best ten features presents an improvement in the classification accuracy as compared with the work in [108], [124] and [41–45] using the NSL-KDD dataset. Following the selection of the best features using FCBFS, accuracy, precision, recall, and F-score are obtained with ML approaches, including DT, RF, NB, ET, and XG-Boost. The findings depict how the classification algorithms may be used in conjunction with the FS approach to increase the system's accuracy and efficiency are shown in Fig. 5.4.

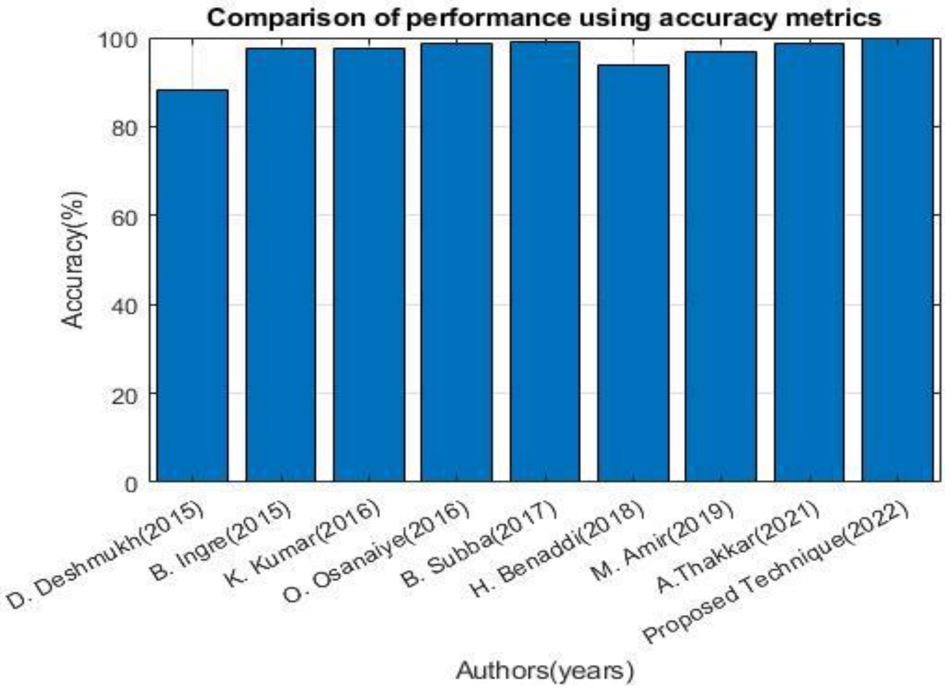


Fig. 5.4. Accuracy comparison with existing methods

The goal of implementing feature engineering is to raise the classification accuracy of the system by choosing pertinent and significant characteristics from the feature collection. The experimental results shown in quantitative and qualitative analysis depict that the combination of the XG-Boost classifier with the proposed FCBFS process yields a comparatively better result with an accuracy of 99.84%. The results showed that our strategy, which uses 10 features, has an improvement in accuracy by 0.17% when compared with [117], 0.71% with [116], 0.18% with [108], and 0.24% with [124].

5.6 Conclusion and Future Scope

In this research, we proposed FCBFS in conjunction with DT, RF, NB, ET, and XG-Boost to pre-process datasets in a heterogeneous cluster-based WSN for IoT applications. The suggested model is trained many times to achieve reproducible results using various classifiers, and the inferences are contrasted with those obtained using the current techniques. Evaluation metrics such as accuracy, precision, recall, and F-score have been calculated to gauge the performance and robustness of the proposed research work. A classic accuracy score of 99.84% is achieved in the case of the XG-Boost classifier, wherein the best ten obtained features were selected after applying the proposed FCBFS method. The proposed technique, which has ten features, outperforms the existing techniques in the literature for NSL-KDD dataset. The obtained findings also show that faster correlation and other classifiers attained higher accuracy, but XG-Boost performed better overall. The proposed technique of the XG-Boost classifier with the FCBFS technique has enhanced the performance and accuracy of WSNs and IoT applications.

In the future, first, we intend to continue our research and see how our strategy affects homogenous WSNs. One of the main issues with the NSL-KDD dataset is the computation time that can be minimized, in our opinion, by using optimization, ensembled, and deep learning algorithms. Second, integration of supervised learning with unsupervised learning, in our opinion, is the best course of action to overcome the overfitting and false positives challenging issues.

Chapter Related Publication

- **Rajiv Yadav,** Indu Sreedevi, Daya Gupta, "Augmentation in Performance and Security of WSNs for IoT Applications using Feature Selection and Classification Techniques," *Alexandria Engineering Journal* 2022, vol. 65, pp. 461-473, **SCIE, (IF: 6.626),** (Accepted & Published), DOI: 10.1016/j.aej.2022.10.033

Chapter 6

Conclusions & Future Directions

In this thesis, we addressed the issues of energy efficiency in WSN, load balanced clustering and routing in WSN and security enhancement in a WSN. The proposed algorithms are scalable and platform independent that is they can be executed on any operating systems. The thesis uses nature-inspired algorithms (GA, PSO, ACO, GWO, and BOA) improve the network lifetime while simultaneously reducing energy consumption and processing delay in order to respond to environmental factors that are helpful for the targeted situation.

Tuning of desirable parameters using hybridization algorithms (GA-PSO, ACO-PSO and PSO-GWO) are also in order to maximize the performance and security of WSNs. These developed frameworks have been compared with state-of-the-art techniques and have given significant results for different performance metrics for all the techniques being compared.

This thesis also addresses a feature selection method (FCBFS) in combination with different classification techniques (DT, RF, NB, ET and XG-Boost) for improving the performance matrices of WSNs. Evaluation metrics such as accuracy, precision, recall, and F1-score have been calculated to gauge the performance and robustness of the proposed work. The proposed technique for XG-Boost classifier with the FCBFS method has enhanced the performance and accuracy of WSNs for IoT applications.

6.1 Summary of the Work Done in the Thesis

This thesis provides a comprehensive review of bio-inspired hybrid optimization algorithms for energy-efficient WSNs. Various advanced techniques in bio-inspired optimization algorithms have been proposed till now, to solve the problem domains in WSN such as data aggregation, sensor location, and routing and coverage area. We have aimed to discuss and compare various newly adopted, hybrid, and conventional methodologies for establishing a robust energy-efficient WSN wherein parameters like packet loss, energy, throughput, delay, and overhead have been utilized. Various open issues and challenges in WSN development using bio-inspired optimization techniques such as network stability, network dynamic character, secure transmission lines, methods to improve QoS, etc., have also been addressed. Hybrid algorithms have better global optimal solutions, search capability, and robustness than other existing algorithms.

For global optimization problems, a variable sensor modality improved butterfly optimization (IBOA) technique is given in this paper. The suggested technique employs a dynamic and adaptive strategy to change the sensor modality, which in the conventional butterfly optimization algorithm was set to a constant value. The butterflies' seeking skills were improved by the varying value of sensor modality. The results showed that the butterflies utilize their knowledge more effectively in the suggested algorithm to execute exploration and exploitation more efficiently than in basic BOA. The comparison is done based on given parameters with existing protocols like LEACH, PSO and GWO.

In this thesis, an FCBFS method in conjunction with DT, RF, NB, ET, and XG-Boost has been proposed to pre-process datasets in a heterogeneous cluster-based WSN for IoT applications. The suggested model is trained many times to achieve reproducible results using various classifiers, and the inferences are contrasted with those obtained using the current techniques. Evaluation metrics such as accuracy, precision, recall, and F1-Score have been calculated to gauge the performance and robustness of the proposed research work. A classic accuracy score of 99.84% is achieved in the case of the XG-Boost classifier, wherein the best ten obtained features were selected after applying the proposed FCBFS. The proposed technique, which has ten features, outperforms the existing techniques in the literature for the NSL-KDD dataset. The obtained findings also show that faster correlation and other classifiers attained higher accuracy, but XG-Boost performed better overall. The proposed technique of the XG-Boost classifier with the FCBFS technique has enhanced the performance and accuracy of WSNs and their IoT applications.

6.2 Future Directions

- The energy efficient algorithms presented in chapter 3 can be enhanced for minimizing the energy consumption of complex IoT and IoE structures. An IoT network consists of large number of heterogeneous nodes, therefore energy conservation and security from different type of attacks is important for its long lasting operation.

- Different unconstrained issues studied in chapter 4; however, it will be fascinating to observe how well the modified butterfly optimization approach performs on restricted problems in upcoming era. Also butterflies utilize their knowledge more effectively to execute exploration and exploitation.

- An FCBFS method in combination with ML classifiers (DT, RF, NB, ET, and XG-Boost) has been proposed chapter 5, so first, we intend to continue our research and see how our strategy affects homogenous WSNs. One of the main issues with the NSL-KDD dataset is the computation time that can be minimized by using optimization and deep learning algorithms. The ensemble approach may be employed for training, and the nature-inspired hybrid algorithms can be utilized for optimization. Second, integration of supervised learning with unsupervised learning, in our opinion, is the best course of action to overcome the overfitting and false positives challenging issues. Third, the performance of IDS models can be enhanced by applying stack-based feature selection techniques.

Publications Related to the Thesis

Papers Published/Accepted for Publication

Journal Papers

1. **Rajiv Yadav,** Indu Sreedevi, Daya Gupta, "Bio-Inspired Hybrid Optimization Algorithms for Energy Efficient Wireless Sensor Networks: A Comprehensive Review," *Electronics*_2022, 11, no. 10, 1545. **SCIE, (IF: 2.690),** (Accepted & Published), DOI: 10.3390/electronics11101545.

2. **Rajiv Yadav,** Indu Sreedevi, Daya Gupta, "Augmentation in Performance and Security of WSNs for IoT Applications using Feature Selection and Classification Techniques," *Alexandria Engineering Journal*_2022, vol. 65, pp. 461-473, **SCIE, (IF: 6.626),** (Accepted & Published), DOI: 10.1016/j.aej.2022.10.033

3. **Rajiv Yadav,** Indu Sreedevi, Daya Gupta, "Security Enhancement for IoT Applications using a Hybrid Nature-inspired Technique," *Wireless Personal Communications,* **SCIE, (IF: 2.017),** (Communicated).

Conference Papers

1. **Rajiv Yadav,** Indu Sreedevi, Daya Gupta, "Review of Evolutionary Algorithms for Energy Efficient and Secure Wireless Sensor Networks," International Conference on Cyber Security and Digital Forensics, pp. 597-608. **Springer, (Scopus Indexed),** Singapore, 2022.DOI: 10.1007/978-981-16-3961-6_49

2. **Rajiv Yadav,** Indu Sreedevi, Daya Gupta, "Performance Analysis of Energy-Efficient Cluster-based Routing Protocols with an Improved Bio-inspired Algorithm in WSNs," International Conference on Data Analytics and Computing, **Springer, (Scopus Indexed),** China, 2022.

Bibliography

[1] A. Shahraki, A. Taherkordi, Ø. Haugen, and F. Eliassen, "Clustering objectives in wireless sensor networks: A survey and research direction analysis," *Comput. Networks*, vol. 180, no. June, 2020, doi: 10.1016/j.comnet.2020.107376.

[2] X. Fu, Y. Yang, and O. Postolache, "Sustainable multipath routing protocol for multi-sink wireless sensor networks in harsh environments," *IEEE Trans. Sustain. Comput.*, vol. 6, no. 1, pp. 168–181, 2021, doi: 10.1109/TSUSC.2020.2976096.

[3] Z. Fei, B. Li, S. Yang, C. Xing, H. Chen, and L. Hanzo, "A Survey of Multi-Objective Optimization in Wireless Sensor Networks: Metrics, Algorithms, and Open Problems," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 1, pp. 550–586, 2017, doi: 10.1109/COMST.2016.2610578.

[4] P. Nayak, G. K. Swetha, S. Gupta, and K. Madhavi, "Routing in wireless sensor networks using machine learning techniques: Challenges and opportunities," *Meas. J. Int. Meas. Confed.*, vol. 178, no. December 2020, p. 108974, 2021, doi: 10.1016/j.measurement.2021.108974.

[5] D. Kandris, C. Nakas, D. Vomvas, and G. Koulouras, "Applications of wireless sensor networks: An up-to-date survey," *Appl. Syst. Innov.*, vol. 3, no. 1, pp. 1–24, 2020, doi: 10.3390/asi3010014.

[6] S. El Khediri, "Wireless sensor networks: a survey, categorization, main issues, and future orientations for clustering protocols," *Computing*, vol. 104, no. 8, pp. 1775–1837, 2022, doi: 10.1007/s00607-022-01071-8.

[7] M. S. Bensaleh, R. Saida, Y. H. Kacem, and M. Abid, "Wireless Sensor Network Design Methodologies: A Survey," *J. Sensors*, vol. 2020, 2020, doi: 10.1155/2020/9592836.

[8] S. Kundu, S. Das, A. V. Vasilakos, and S. Biswas, "A modified differential evolution-based combined routing and sleep scheduling scheme for lifetime maximization of wireless sensor networks," *Soft Comput.*, vol. 19, no. 3, pp. 637–659, 2015, doi: 10.1007/s00500-014-1286-9.

[9] A. Kaushik, S. Indu, and D. Gupta, "A novel load balanced energy conservation approach in WSN using biogeography based optimization," *AIP Conf. Proc.*, vol. 1884, pp. 1–4, 2017, doi: 10.1063/1.5002507.

[10] M. S. Yousefpoor, E. Yousefpoor, H. Barati, A. Barati, A. Movaghar, and M. Hosseinzadeh, "Secure data aggregation methods and countermeasures against various attacks in wireless sensor networks: A comprehensive review," *J. Netw. Comput. Appl.*, vol. 190, no. May, p. 103118, 2021, doi: 10.1016/j.jnca.2021.103118.

[11] T. M. Behera, S. K. Mohapatra, U. C. Samal, M. S. Khan, M. Daneshmand, and A. H. Gandomi, "Residual energy-based cluster-head selection in WSNs for IoT application," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5132–5139, 2019, doi: 10.1109/JIOT.2019.2897119.

[12] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, and B. Sikdar, "A Survey on IoT Security: Application Areas, Security Threats, and Solution Architectures," *IEEE Access*, vol. 7, pp. 82721–82743, 2019, doi: 10.1109/ACCESS.2019.2924045.

[13] M. Elshrkawey, S. M. Elsherif, and M. Elsayed Wahed, "An Enhancement Approach for Reducing the Energy Consumption in Wireless Sensor Networks," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 30, no. 2, pp. 259–267, 2018, doi: 10.1016/j.jksuci.2017.04.002.

[14] L. Nguyen and H. T. Nguyen, "Mobility based network lifetime in wireless sensor networks: A review," *Comput. Networks*, vol. 174, no. November 2019, p. 107236, 2020, doi: 10.1016/j.comnet.2020.107236.

[15] S. Sudha, B. Manimegalai, and P. Thirumoorthy, "A study on routing approach for in-network aggregation in wireless sensor networks," *2014 Int. Conf. Comput. Commun. Informatics Ushering Technol. Tomorrow, Today, ICCCI 2014*, pp. 6–9, 2014, doi: 10.1109/ICCCI.2014.6921834.

[16] H. Oh and K. Chae, "An Energy-Efficient sensor routing with low latency, scalability for Smart Home Networks," *Int. J. Smart Home*, vol. 1, no. 2, pp. 71–82, 2007 doi: 10.1109/MUE.2007.75.

[17] M. Farsi, M. A. Elhosseini, M. Badawy, H. Arafat Ali, and H. Zain Eldin, "Deployment techniques in wireless sensor networks, coverage and connectivity: A survey," *IEEE Access*, vol. 7, pp. 28940–28954, 2019, doi: 10.1109/ACCESS.2019.2902072.

[18] Z. Huanan, X. Suping, and W. Jiannan, "Security and application of wireless sensor network," *Procedia Comput. Sci.*, vol. 183, pp. 486–492, 2021, doi: 10.1016/j.procs.2021.02.088.

[19] K. Jaiswal and V. Anand, "EOMR: An Energy-Efficient Optimal Multi-path Routing Protocol to Improve QoS in Wireless Sensor Network for IoT Applications," *Wirel. Pers. Commun.*, vol. 111, no. 4, pp. 2493–2515, 2020, doi: 10.1007/s11277-019-07000-x.

[20] V. T. Pham, T. N. Nguyen, B. H. Liu, and T. Lin, "Minimizing latency for multiple-type data aggregation in wireless sensor networks," *IEEE Wirel. Commun. Netw. Conf. WCNC*, vol. 2021-March, 2021, doi: 10.1109/WCNC49053.2021.9417309.

[21] K. Skiadopoulos *et al.*, "Synchronization of data measurements in wireless sensor networks for IoT applications," *Ad Hoc Networks*, vol. 89, pp. 47–57, 2019, doi: 10.1016/j.adhoc.2019.03.002.

[22] N. Sharma and V. Gupta, "Meta-heuristic based optimization of WSNs Localisation Problem-a Survey," *Procedia Comput. Sci.*, vol. 173, no. 2019, pp. 36–45, 2020, doi: 10.1016/j.procs.2020.06.006.

[23] A. Singh, S. Sharma, and J. Singh, "Nature-inspired algorithms for Wireless Sensor Networks: A comprehensive survey," *Comput. Sci. Rev.*, vol. 39, p. 100342, 2021, doi: 10.1016/j.cosrev.2020.100342.

[24] K.-L. Du and M. N. S. Swamy, "Search and Optimization by Metaheuristics," *Tech. and Algo. Inspired by Nature,* vol. 1, pp. 1-10, 2016, 10.1007/978-3-319-41192-7.

[25] J. H. Holland, "Genetic Algorithms - Computer programs that 'evolve' in ways that resemble natural selection can solve complex problems even their creators do not fully understand," *Sci. Am.*, pp. 66–72, 1992, doi: 10.1007/0-387-28356-0_4.

[26] J. Rejinaparvin and C. Vasanthanayaki, "Particle swarm optimization-based clustering by preventing residual nodes in wireless sensor networks," *IEEE Sens. J.*, vol. 15, no. 8, pp. 4264–4274, 2015, doi: 10.1109/JSEN.2015.2416208.

[27] Y. L. Wang, M. Song, Y. F. Wei, Y. H. Wang, and X. J. Wang, "Improved ant colony-based multi-constrained QoS energy-saving routing and throughput optimization in wireless Ad-hoc networks," *J. China Univ. Posts Telecommun.*, vol. 21, no. 1, pp. 43–53, 2014, doi: 10.1016/S1005-8885(14)60267-3.

[28] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[29] A. Kaushik, S. Indu, and D. Gupta, "A Grey Wolf Optimization Approach for Improving the Performance of Wireless Sensor Networks," *Wirel. Pers. Commun.*, vol. 106, no. 3, pp. 1429–1449, 2019, doi: 10.1007/s11277-019-06223-2.

[30] X. S. Yang, "Nature-inspired optimization algorithms: Challenges and open problems," *J. Comput. Sci.*, vol. 46, pp. 1–15, 2020, doi: 10.1016/j.jocs.2020.101104.

[31] M. Shokouhifar and A. Jalali, "A new evolutionary based application specific routing protocol for clustered wireless sensor networks," *AEU - Int. J. Electron. Commun.*, vol. 69, no. 1, pp. 432–441, 2015, doi: 10.1016/j.aeue.2014.10.023.

[32] P. Srideviponmalar, V. Jawahar Senthil Kumar, and R. Harikrishnan, *Hybrid genetic algorithm–differential evolution approach for localization in WSN*, vol. 695. Springer Singapore, 2018, doi: 10.1007/978-981-10-7566-7_27

[33] A. Agnihotri and I. K. Gupta, "A hybrid PSO-GA algorithm for routing in

wireless sensor network," *Proc. 4th IEEE Int. Conf. Recent Adv. Inf. Technol. RAIT 2018*, pp. 1–6, 2018, doi: 10.1109/RAIT.2018.8389082.

[34] R. Rastogi, S. Srivastava, Tarun, M. Singh Manshahia, Varsha, and N. Kumar, "A hybrid optimization approach using PSO and ant colony in wireless sensor network," *Mater. Today Proc.*, no. xxxx, 2021, doi: 10.1016/j.matpr.2021.01.874.

[35] F. A. Şenel, F. Gökçe, A. S. Yüksel, and T. Yiğit, "A novel hybrid PSO–GWO algorithm for optimization problems," *Eng. Comput.*, vol. 35, no. 4, pp. 1359–1373, 2019, doi: 10.1007/s00366-018-0668-5.

[36] S. Prithi and S. Sumathi, "Automata Based Hybrid PSO–GWO Algorithm for Secured Energy Efficient Optimal Routing in Wireless Sensor Network," *Wirel. Pers. Commun.*, vol. 117, no. 2, pp. 545–559, 2021, doi: 10.1007/s11277-020-07882-2.

[37] A. Gambhir, A. Payal, and R. Arya, "Performance analysis of artificial bee colony optimization based clustering protocol in various scenarios of WSN," *Procedia Comput. Sci.*, vol. 132, pp. 183–188, 2018, doi: 10.1016/j.procs.2018.05.184.

[38] O. O. Ogundile and A. S. Alfa, "A survey on an energy-efficient and energy-balanced routing protocol for wireless sensor networks," *Sensors (Switzerland)*, vol. 17, no. 5, pp. 1–51, 2017, doi: 10.3390/s17051084.

[39] M. Abdullah, A. Balamash, A. Al-Shannaq, and S. Almabdy, "Enhanced Intrusion Detection System using Feature Selection Method and Ensemble Learning Algorithms," *Int. J. Comput. Sci. Inf. Secur.*, vol. 16, no. December, pp. 48–55, 2018, doi: 10.14355/ijcsa.2018.0304.02.

[40] O. Osanaiye, O. Ogundile, F. Aina, and A. Periola, "Feature selection for intrusion detection system in a cluster-based heterogeneous wireless sensor network," *Facta Univ. - Ser. Electron. Energ.*, vol. 32, no. 2, pp. 315–330, 2019, doi: 10.2298/fuee1902315o.

[41] D. E. Boubiche, S. Athmani, S. Boubiche, and H. Toral-Cruz, "Cybersecurity Issues in Wireless Sensor Networks: Current Challenges and Solutions," *Wirel. Pers. Commun.*,vol. 117, no. 1. Springer US, 2021,doi: 10.1007/s11277-020-07213-5.

[42] C. C. Lee, "Security and privacy in wireless sensor networks: Advances and challenges," *Sensors (Switzerland)*, vol. 20, no. 3, pp. 10–12, 2020, doi: 10.3390/s20030744.

[43] R. Yugha and S. Chithra, "A survey on technologies and security protocols: Reference for future generation IoT," *J. Netw. Comput. Appl.*, vol. 169, no. August, p. 102763, 2020, doi: 10.1016/j.jnca.2020.102763.

[44] A. Wang, D. Yang, and D. Sun, "A clustering algorithm based on energy information and cluster heads expectation for wireless sensor networks,"

Comput. Electr. Eng., vol. 38, no. 3, pp. 662–671, 2012, doi: 10.1016/j.compeleceng.2011.11.017.

[45] B. Manzoor *et al.*, "Q-LEACH: A new routing protocol for WSNs," *Procedia Comput. Sci.*, vol. 19, pp. 926–931, 2013, doi: 10.1016/j.procs.2013.06.127.

[46] B. A. Attea and E. A. Khalil, "A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks," *Appl. Soft Comput. J.*, vol. 12, no. 7, pp. 1950–1957, 2012, doi: 10.1016/j.asoc.2011.04.007.

[47] M. Barcelo, A. Correa, J. Lopez Vicario, and A. Morell, "Cooperative interaction among multiple RPL instances in wireless sensor networks," *Comput. Commun.*, vol. 81, pp. 61–71, 2016, doi: 10.1016/j.comcom.2015.12.008.

[48] J. R. Srivastava and T. S. B. Sudarshan, "A genetic fuzzy system based optimized zone based energy efficient routing protocol for mobile sensor networks (OZEEP)," *Appl. Soft Comput. J.*, vol. 37, pp. 863–886, 2015, doi: 10.1016/j.asoc.2015.09.025.

[49] C. W. Chen and C. C. Weng, "A power efficiency routing and maintenance protocol in wireless multi-hop networks," *J. Syst. Softw.*, vol. 85, no. 1, pp. 62–76, 2012, doi: 10.1016/j.jss.2011.07.012.

[50] T. Hayes and F. H. Ali, "Proactive Highly Ambulatory Sensor Routing (PHASeR) protocol for mobile wireless sensor networks," *Pervasive Mob. Comput.*, vol. 21, pp. 47–61, 2015, doi: 10.1016/j.pmcj.2015.04.005.

[51] S. Singh, S. Chand, and B. Kumar, "Energy Efficient Clustering Protocol Using Fuzzy Logic for Heterogeneous WSNs," *Wirel. Pers. Commun.*, vol. 86, no. 2, pp. 451–475, 2016, doi: 10.1007/s11277-015-2939-4.

[52] H. Kumar, H. Arora, and R. K. Singla, "Energy-aware fisheye routing (EA-FSR) algorithm for wireless mobile sensor networks," *Egypt. Informatics J.*, vol. 14, no. 3, pp. 235–238, 2013, doi: 10.1016/j.eij.2013.07.002.

[53] J. Wang, Y. Cao, B. Li, H. jin Kim, and S. Lee, "Particle swarm optimization based clustering algorithm with mobile sink for WSNs," *Futur. Gener. Comput. Syst.*, vol. 76, pp. 452–457, 2017, doi: 10.1016/j.future.2016.08.004.

[54] A. O. Helmy, S. Ahmed, and A. E. Hassenian, "Artificial fish swarm algorithm for energy-efficient routing technique," *Adv. Intell. Syst. Comput.*, vol. 322, pp. 509–519, 2015, doi: 10.1007/978-3-319-11313-5_45.

[55] M. Liu, S. Xu, and S. Sun, "An agent-assisted QoS-based routing algorithm for wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 35, no. 1, pp. 29–36, 2012, doi: 10.1016/j.jnca.2011.03.031.

[56] P. Kuila and P. K. Jana, "Energy efficient clustering and routing algorithms for wireless sensor networks: Particle swarm optimization approach," *Eng. Appl. Artif. Intell.*, vol. 33, pp. 127–140, 2014, doi:

10.1016/j.engappai.2014.04.009.

[57] P. C. S. Rao, P. K. Jana, and H. Banka, “A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks,” *Wirel. Networks*, vol. 23, no. 7, pp. 2005–2020, 2017, doi: 10.1007/s11276-016-1270-7.

[58] T. Shankar, S. Shanmugavel, and A. Rajesh, “Hybrid HSA and PSO algorithm for energy efficient cluster head selection in wireless sensor networks,” *Swarm Evol. Comput.*, vol. 30, pp. 1–10, 2016, doi: 10.1016/j.swevo.2016.03.003.

[59] Z. M. Zahedi, R. Akbari, M. Shokouhifar, F. Safaei, and A. Jalali, “Swarm intelligence based fuzzy routing protocol for clustered wireless sensor networks,” *Expert Syst. Appl.*, vol. 55, pp. 313–328, 2016, doi: 10.1016/j.eswa.2016.02.016.

[60] D. gan Zhang, X. Wang, X. dong Song, T. Zhang, and Y. nan Zhu, “A new clustering routing method based on PECE for WSN,” *Eurasip J. Wirel. Commun. Netw.*, vol. 2015, no. 1, 2015, doi: 10.1186/s13638-015-0399-x.

[61] B. Zeng and Y. Dong, “An improved harmony search based energy-efficient routing algorithm for wireless sensor networks,” *Appl. Soft Comput. J.*, vol. 41, pp. 135–147, 2016, doi: 10.1016/j.asoc.2015.12.028.

[62] G. S. Brar, S. Rani, V. Chopra, R. Malhotra, H. Song, and S. H. Ahmed, “Energy efficient direction-based PDORP routing protocol for WSN,” *IEEE Access*, vol. 4, pp. 3182–3194, 2016, doi: 10.1109/ACCESS.2016.2576475.

[63] R. R. Sahoo, M. Singh, B. M. Sahoo, K. Majumder, S. Ray, and S. K. Sarkar, “A Light Weight Trust based Secure and Energy Efficient Clustering in Wireless Sensor Network: Honey Bee Mating Intelligence Approach,” *Procedia Technol.*, vol. 10, pp. 515–523, 2013, doi: 10.1016/j.protcy.2013.12.390.

[64] A. Adamou Abba Ari, B. Omer Yenke, N. Labraoui, I. Damakoa, and A. Gueroui, “A power efficient cluster-based routing algorithm for wireless sensor networks: Honeybees swarm intelligence based approach,” *J. Netw. Comput. Appl.*, vol. 69, pp. 77–97, 2016, doi: 10.1016/j.jnca.2016.04.020.

[65] S. Gajjar, M. Sarkar, and K. Dasgupta, “FAMACROW: Fuzzy and ant colony optimization based combined mac, routing, and unequal clustering cross-layer protocol for wireless sensor networks,” *Appl. Soft Comput. J.*, vol. 43, pp. 235–247, 2016, doi: 10.1016/j.asoc.2016.02.019.

[66] P. S. Mann and S. Singh, “Energy-Efficient Hierarchical Routing for Wireless Sensor Networks: A Swarm Intelligence Approach,” *Wirel. Pers. Commun.*, vol. 92, no. 2, pp. 785–805, 2017, doi: 10.1007/s11277-016-3577-1.

[67] M. Saleem, I. Ullah, and M. Farooq, “BeeSensor: An energy-efficient and scalable routing protocol for wireless sensor networks,” *Inf. Sci. (Ny).*, vol.

200, pp. 38–56, 2012, doi: 10.1016/j.ins.2012.02.024.

[68] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: Optimization by a colony of cooperating agents," *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*, vol. 26, no. 1, pp. 29–41, 1996, doi: 10.1109/3477.484436.

[69] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm," *J. Glob. Optim.*, vol. 39, no. 3, pp. 459–471, 2007, doi: 10.1007/s10898-007-9149-x.

[70] D. Simon and S. Member, "Biogeography-Based Optimization," vol. 12, no. 6, pp. 702–713, 2008, doi: 10.1109/TEVC.2008.919004.

[71] H. Shah-Hosseini, "The intelligent water drops algorithm: A nature-inspired swarm-based optimization algorithm," *Int. J. Bio-Inspired Comput.*, vol. 1, no. 1–2, pp. 71–79, 2009, doi: 10.1504/IJBIC.2009.022775.

[72] E. Rashedi, H. Nezamabadi-pour, and S. Saryazdi, "GSA: A Gravitational Search Algorithm," *Inf. Sci. (Ny).*, vol. 179, no. 13, pp. 2232–2248, 2009, doi: 10.1016/j.ins.2009.03.004.

[73] X. S. Yang, "A new metaheuristic Bat-inspired Algorithm," *Stud. Comput. Intell.*, vol. 284, pp. 65–74, 2010, doi: 10.1007/978-3-642-12538-6_6.

[74] A. Kaveh and S. Talatahari, "A novel heuristic optimization method: Charged system search," *Acta Mech.*, vol. 213, no. 3–4, pp. 267–289, 2010, doi: 10.1007/s00707-009-0270-4.

[75] M. Clerc, "Particle Swarm Optimization," *Part. Swarm Optim.*, pp. 1942–1948, 2010, doi: 10.1002/9780470612163.

[76] S. Goel, A. Sharma, and P. Bedi, "Cuckoo search clustering algorithm: A novel strategy of biomimicry," *Proc. 2011 World Congr. Inf. Commun. Technol. WICT 2011*, pp. 916–921, 2011, doi: 10.1109/WICT.2011.6141370.

[77] X. S. Yang, "Flower pollination algorithm for global optimization," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7445 LNCS, pp. 240–249, 2012, doi: 10.1007/978-3-642-32894-7_27.

[78] A. Sabry Eesa, A. Mohsin, A. Brifcani, and Z. Orman, "Cuttlefish Algorithm – A Novel Bio-Inspired Optimization Algorithm," *Int. J. Sci. Eng. Res.*, vol. 4, no. 9, pp. 1978–1986, 2013, [Online]. Available: http://www.ijser.org.

[79] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, 2016, doi: 10.1016/j.advengsoft.2016.01.008.

[80] S. Gao and Y. Wen, "An Improved Artificial Fish Swarm Algorithm and its Application," *Proc. - 17th IEEE/ACIS Int. Conf. Comput. Inf. Sci. ICIS 2018*, pp. 649–652, 2018, doi: 10.1109/ICIS.2018.8466458.

[81] Y. Y. Hao, G. L. Zhang, and B. Xiong, “An Improved Glowworm Swarm Optimization Algorithm,” *Proc. - Int. Conf. Mach. Learn. Cybern.*, vol. 1, pp. 155–160, 2018, doi: 10.1109/ICMLC.2018.8526982.

[82] S. Su and S. Zhao, “A hierarchical hybrid of genetic algorithm and particle swarm optimization for distributed clustering in large-scale wireless sensor networks,” *J. Ambient Intell. Humaniz. Comput.*, vol. 0, no. 0, pp. 1–11, 2017, doi: 10.1007/s12652-017-0619-9.

[83] J. Kapoor, S. Goel, K. K. Sharma, and M. Tech, “A GA-BFO Swarm Based CH Selection Approach For Enhancing Efficiency In Clustering Protocols,” vol. 2, no. 1, pp. 16–20, 2017.

[84] B. Farnad, A. Jafarian, and D. Baleanu, “A new hybrid algorithm for continuous optimization problem,” *Appl. Math. Model.*, vol. 55, pp. 652–673, 2018, doi: 10.1016/j.apm.2017.10.001.

[85] S. Potthuri, T. Shankar, and A. Rajesh, “Lifetime Improvement in Wireless Sensor Networks using Hybrid Differential Evolution and Simulated Annealing (DESA),” *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 655–663, 2018, doi: 10.1016/j.asej.2016.03.004.

[86] B. Pitchaimanickam and G. Murugaboopathi, “A hybrid firefly algorithm with particle swarm optimization for energy efficient optimal cluster head selection in wireless sensor networks,” *Neural Comput. Appl.*, vol. 32, no. 12, pp. 7709–7723, 2020, doi: 10.1007/s00521-019-04441-0.

[87] L. Nagarajan and S. Thangavelu, “Hybrid grey wolf sunflower optimisation algorithm for energy-efficient cluster head selection in wireless sensor networks for lifetime enhancement,” *IET Commun.*, vol. 15, no. 3, pp. 384–396, 2021, doi: 10.1049/cmu2.12072.

[88] K. Q. Yan, S. C. Wang, S. S. Wang, and C. W. Liu, “Hybrid Intrusion Detection System for enhancing the security of a cluster-based Wireless Sensor Network,” *Proc. - 2010 3rd IEEE Int. Conf. Comput. Sci. Inf. Technol. ICCSIT 2010*, vol. 1, pp. 114–118, 2010, doi: 10.1109/ICCSIT.2010.5563886.

[89] A. I. Saleh, F. M. Talaat, and L. M. Labib, “A hybrid intrusion detection system (HIDS) based on prioritized k-nearest neighbors and optimized SVM classifiers,” *Artif. Intell. Rev.*, vol. 51, no. 3, pp. 403–443, 2019, doi: 10.1007/s10462-017-9567-1.

[90] S. W. Lin, K. C. Ying, C. Y. Lee, and Z. J. Lee, “An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection,” *Appl. Soft Comput. J.*, vol. 12, no. 10, pp. 3285–3290, 2012, doi: 10.1016/j.asoc.2012.05.004.

[91] A. M. Ahmed, A. Rizaner, and A. H. Ulusoy, “A Decision Tree Algorithm Combined with Linear Regression for Data Classification,” *2018 Int. Conf. Comput. Control. Electr. Electron. Eng. ICCCEEE 2018*, 2018, doi: 10.1109/ICCCEEE.2018.8515759.

[92] S. Bhattacharya and S. Selvakumar, “Multi-measure multi-weight ranking approach for the identification of the network features for the detection of DoS and probe attacks,” *Comput. J.*, vol. 59, no. 6, pp. 923–943, 2016, doi: 10.1093/comjnl/bxv078.

[93] Ü. Çavuşoğlu, “(NSL-KDD &CfsSubsetEval & WrapperSubsetEval)A new hybrid approach for intrusion detection using machine learning methods,” *Appl. Intell.*, vol. 49, no. 7, pp. 2735–2761, 2019, doi: 10.1007/s10489-018-01408-x.

[94] K. Medhat, R. A. Ramadan, and I. Talkhan, “Distributed Intrusion Detection System for Wireless Sensor Networks,” *Proc. - NGMAST 2015 9th Int. Conf. Next Gener. Mob. Appl. Serv. Technol.*, pp. 234–239, 2016, doi: 10.1109/NGMAST.2015.29.

[95] M. Safaldin, M. Otair, and L. Abualigah, “Improved binary gray wolf optimizer and SVM for intrusion detection system in wireless sensor networks,” *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 2, pp. 1559–1576, 2021, doi: 10.1007/s12652-020-02228-z.

[96] S. Potluri and C. Diedrich, “Accelerated deep neural networks for enhanced Intrusion Detection System,” *IEEE Int. Conf. Emerg. Technol. Fact. Autom. ETFA*, vol. 2016-Novem, pp. 1–8, 2016, doi: 10.1109/ETFA.2016.7733515.

[97] A. Zhou, Z. Li, and Y. Shen, “Anomaly detection of CAN bus messages using a deep neural network for autonomous vehicles,” *Appl. Sci.*, vol. 9, no. 15, 2019, doi: 10.3390/app9153174.

[98] I. Almomani, B. Al-Kasasbeh, and M. Al-Akhras, “WSN-DS: A Dataset for Intrusion Detection Systems in Wireless Sensor Networks,” *J. Sensors*, vol. 2016, 2016, doi: 10.1155/2016/4731953.

[99] A. I. Al-Issa, M. Al-Akhras, M. S. Alsahli, and M. Alawairdhi, “Using machine learning to detect dos attacks in wireless sensor networks,” *2019 IEEE Jordan Int. Jt. Conf. Electr. Eng. Inf. Technol. JEEIT 2019 - Proc.*, pp. 107–112, 2019, doi: 10.1109/JEEIT.2019.8717400.

[100] K. Alrawashdeh and C. Purdy, “Toward an online anomaly intrusion detection system based on deep learning,” *Proc. - 2016 15th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2016*, pp. 195–200, 2017, doi: 10.1109/ICMLA.2016.167.

[101] B. Ghojogh, A. Ghodsi, F. Karray, and M. Crowley, “Restricted Boltzmann Machine and Deep Belief Network: Tutorial and Survey,” 2021, [Online]. Available: http://arxiv.org/abs/2107.12521.

[102] R. Doshi, N. Apthorpe, and N. Feamster, “Machine learning DDoS detection for consumer internet of things devices,” *Proc. - 2018 IEEE Symp. Secur. Priv. Work. SPW 2018*, no. Ml, pp. 29–35, 2018, doi: 10.1109/SPW.2018.00013.

[103] Y. Cheng, Y. Xu, H. Zhong, and Y. Liu, "Leveraging Semisupervised Hierarchical Stacking Temporal Convolutional Network for Anomaly Detection in IoT Communication," *IEEE Internet Things J.*, vol. 8, no. 1, pp. 144–155, 2021, doi: 10.1109/JIOT.2020.3000771.

[104] M. Zekri, S. El Kafhali, N. Aboutabit, and Y. Saadi, "DDoS attack detection using machine learning techniques in cloud computing environments," *Proc. 2017 Int. Conf. Cloud Comput. Technol. Appl. CloudTech 2017*, vol. 2018-Janua, no. October, pp. 1–7, 2018, doi: 10.1109/CloudTech.2017.8284731.

[105] S. Sumathi, R. Rajesh, and N. Karthikeyan, "DDoS Attack Detection Using Hybrid Machine Learning Based IDS Models," *J. Sci. Ind. Res. (India).*, vol. 81, no. 3, pp. 276–286, 2022, doi: 10.56042/jsir.v81i03.58451.

[106] A. Khamparia, S. Pande, D. Gupta, A. Khanna, and A. K. Sangaiah, "Multi-level framework for anomaly detection in social networking," *Libr. Hi Tech*, vol. 38, no. 2, pp. 350–366, 2020, doi: 10.1108/LHT-01-2019-0023.

[107] M. Aldwairi and L. Tawalbeh, "Security techniques for intelligent spam sensing and anomaly detection in online social platforms," *Int. J. Electr. Comput. Eng.*, vol. 10, no. 1, pp. 275–287, 2020, doi: 10.11591/ijece.v10i1.pp275-287.

[108] M. Aamir and S. M. Ali Zaidi, "Clustering based semi-supervised machine learning for DDoS attack classification," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 33, no. 4, pp. 436–446, 2021, doi: 10.1016/j.jksuci.2019.02.003.

[109] L. Su, Y. Yao, N. Li, J. Liu, Z. Lu, and B. Liu, "Hierarchical Clustering Based Network Traffic Data Reduction for Improving Suspicious Flow Detection," *Proc. - 17th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. 12th IEEE Int. Conf. Big Data Sci. Eng. Trust. 2018*, pp. 744–753, 2018, doi: 10.1109/TrustCom/BigDataSE.2018.00108.

[110] I. S. Kocher, "A Systematical Roadmap on Various Security Vulnerabilities and Countermeasures in Routing Algorithms upon WSNs," *Acad. J. Nawroz Univ.*, vol. 10, no. 4, pp. 1–13, 2021, doi: 10.25007/ajnu.v10n4a957.

[111] I. S. Kocher, "An experimental simulation of addressing auto-configuration issues for wireless sensor networks," *Comput. Mater. Contin.*, vol. 71, no. 2, pp. 3821–3838, 2022, doi: 10.32604/cmc.2022.023478.

[112] M. R. Norouzian and S. Merati, "Classifying attacks in a network intrusion detection system based on artificial neural networks," *Int. Conf. Adv. Commun. Technol. ICACT*, pp. 868–873, 2011, doi: 10.1007/s10796-009-9210-z .

[113] S. Mukherjee and N. Sharma, "Intrusion Detection using Naive Bayes Classifier with Feature Reduction," *Procedia Technol.*, vol. 4, pp. 119–128, 2012, doi: 10.1016/j.protcy.2012.05.017.

[114] H. Chae, B. Jo, S. Choi, and T. Park, "Feature Selection for Intrusion

Detection using NSL-KDD," *Recent Adv. Comput. Sci. 20132*, pp. 184–187, 2013, doi: 10.1109/RACS.2013.6380910.

[115] S. T. Ikram and A. K. Cherukuri, "Improving accuracy of intrusion detection model using PCA and optimized SVM," *J. Comput. Inf. Technol.*, vol. 24, no. 2, pp. 133–148, 2016, doi: 10.20532/cit.2016.1002701.

[116] B. Subba, S. Biswas, and S. Karmakar, "Enhancing performance of anomaly based intrusion detection systems through dimensionality reduction using principal component analysis," *2016 IEEE Int. Conf. Adv. Networks Telecommun. Syst. ANTS 2016*, 2017, doi: 10.1109/ANTS.2016.7947776.

[117] O. Osanaiye, H. Cai, K. K. R. Choo, A. Dehghantanha, Z. Xu, and M. Dlodlo, "Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing," *Eurasip J. Wirel. Commun. Netw.*, vol. 2016, no. 1, 2016, doi: 10.1186/s13638-016-0623-3.

[118] K. Kumar and J. Singh, "Network Intrusion Detection with Feature Selection Techniques using Machine-Learning Algorithms," *Int. J. Comput. Appl.*, vol. 150, no. 12, pp. 1–13, 2016, doi: 10.5120/ijca2016910764.

[119] I. Sumaiya Thaseen and C. Aswani Kumar, "Intrusion detection model using fusion of chi-square feature selection and multi class SVM," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 29, no. 4, pp. 462–472, 2017, doi: 10.1016/j.jksuci.2015.12.004.

[120] S. Meftah, T. Rachidi, and N. Assem, "Network Based Intrusion Detection Using the UNSW-NB15 dataset," *Int. J. Comput. Netw.,* vol. 5, no. 5, 2019, doi: 10.12785/ijcds/080505.

[121] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman, and A. Alazab, "Hybrid intrusion detection system based on the stacking ensemble of C5 decision tree classifier and one class support vector machine," *Electron.*, vol. 9, no. 1, 2020, doi: 10.3390/electronics9010173.

[122] B. S. Bhati and C. S. Rai, "Analysis of Support Vector Machine-based Intrusion Detection Techniques," *Arab. J. Sci. Eng.*, vol. 45, no. 4, pp. 2371–2383, 2020, doi: 10.1007/s13369-019-03970-z.

[123] S. Vignesh and B. R. Kanna, "AWS Infrastructure Automation and Security Prevention Using DevOps, " *Art. Intell. Comp. Eng. Sys.,* vol. 1056. pp 537–549, 2020, doi: 10.1007/978-981-15-0199-9_46.

[124] A. Thakkar and R. Lohiya, "Attack classification using feature selection techniques: a comparative study," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 1, pp. 1249–1266, 2021, doi: 10.1007/s12652-020-02167-9.

[125] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, vol. 2000-Janua, no. c, pp. 1–10, 2000.

[126] I. Gupta, D. Riordan, and S. Sampalli, "Cluster-head election using fuzzy logic for wireless sensor networks," *Proc. 3rd Annu. Commun. Networks Serv. Res. Conf.*, vol. 2005, pp. 255–260, 2005, doi: 10.1109/CNSR.2005.27.

[127] Y. K. Tamandani and M. U. Bokhari, "SEPFL routing protocol based on fuzzy logic control to extend the lifetime and throughput of the wireless sensor network," *Wirel. Networks*, vol. 22, no. 2, pp. 647–653, 2016, doi: 10.1007/s11276-015-0997-x.

[128] M. A. AbdulAlim, Y. C. Wu, and W. Wang, "A fuzzy based clustering protocol for energy-efficient wireless sensor networks," *Adv. Mater. Res.*, vol. 760–762, no. Iccsee, pp. 685–690, 2013, doi: 10.4028/www.scientific.net/AMR.760-762.685.

[129] S. K. Gupta, P. Kuila, and P. K. Jana, "GAR: An energy efficient GA-based routing for wireless sensor networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7753 LNCS, pp. 267–277, 2013, doi: 10.1007/978-3-642-36071-8_21.

[130] C. Blum and M. López-Ibáñez, "Ant colony optimization," *Intell. Syst.*, no. 2, pp. 1461, 2016, doi: 10.4249/scholarpedia.1461.

[131] A. Kaushik, S. Indu, and D. Gupta, "Optimizing and enhancing the lifetime of a wireless sensor network using biogeography based optimization," *Commun. Comput. Inf. Sci.*, vol. 899, pp. 260–272, 2019, doi: 10.1007/978-981-13-2035-4_23.

[132] S. Zhang, S. Xu, W. Zhang, D. Yu, and K. Chen, "A hybrid approach combining an extended BBO algorithm with an intuitionistic fuzzy entropy weight method for QoS-aware manufacturing service supply chain optimization," *Neurocomputing*, vol. 272, pp. 439–452, 2018, doi: 10.1016/j.neucom.2017.07.011.

[133] K. Shukla, M. Verma, and D. Gupta, "Age-Structured Biogeography-based Optimization," *Int. Conf. Intell. Comp. Netw. Syst. (Iciccs),* no. 2, pp. 339–346, 2020, doi: 10.1109/ICICCS48265.2020.9121034.

[134] C. P. Low, C. Fang, J. M. Ng, and Y. H. Ang, "Efficient Load-Balanced Clustering Algorithms for wireless sensor networks," *Comput. Commun.*, vol. 31, no. 4, pp. 750–759, 2008, doi: 10.1016/j.comcom.2007.10.020.

[135] J. He, S. Ji, M. Yan, Y. Pan, and Y. Li, "Load-balanced CDS construction in wireless sensor networks via genetic algorithm," *Int. J. Sens. Networks*, vol. 11, no. 3, pp. 166–178, 2012, doi: 10.1504/IJSNET.2012.046331.

[136] A. Raha, M. Kanti Naskar, A. Paul, A. Chakraborty, and A. Karmakar, "A Genetic Algorithm Inspired Load Balancing Protocol for Congestion Control in Wireless Sensor Networks using Trust Based Routing Framework (GACCTR)," *Int. J. Comput. Netw. Inf. Secur.*, vol. 5, no. 9, pp. 9–20, 2013, doi: 10.5815/ijcnis.2013.09.02.

[137] S. Balaji, E. G. Julie, M. Rajaram, and Y. H. Robinson, "Fuzzy Based Particle Swarm Optimization Routing Technique for Load Balancing in Wireless Sensor Networks," *Int. J. Comput. Inf. Eng.*, vol. 10, no. 7, pp. 1418–1427, 2016, doi: 10.6064/IJCIE.2016.2177802.

[138] X. Liu, T. Qiu, and T. Wang, "Load-Balanced Data Dissemination for Wireless Sensor Networks: A Nature-Inspired Approach," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 9256–9265, 2019, doi: 10.1109/JIOT.2019.2900763.

[139] O. Bozorg-Haddad, "Studies in Computational Intelligence - Advanced Optimization by Nature-Inspired Algorithms," *Stud. Comp. Intell.,* vol. 720, 2018, doi: 10.1018/SCIS/2018.52217.

[140] M. Patel *et al.*, "Security attacks in wireless sensor networks," pp. 536–539, 2013, doi: 10.1109/ISSP.2013.6526929.

[141] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler, "SPINS: Security protocols for sensor networks," *Wirel. Networks*, vol. 8, no. 5, pp. 521–534, 2002, doi: 10.1023/A:1016598314198.

[142] M. Conti, R. Di Pietro, L. Mancini, and A. Mei, "Distributed detection of clone attacks in wireless sensor networks," *IEEE Trans. Dependable Secur. Comput.*, vol. 8, no. 5, pp. 685–698, 2011, doi: 10.1109/TDSC.2010.25.

[143] V. D. Gligor, "Emergent properties in ad-hoc networks: A security perspective," *Proc. 2006 ACM Symp. Information, Comput. Commun. Secur. ASIACCS '06*, vol. 2006, p. 1, 2006, doi: 10.1145/1128817.1128819.

[144] J. Ren, Y. Zhang, K. Zhang, and X. Shen, "Adaptive and channel-aware detection of selective forwarding attacks in wireless sensor networks," *IEEE Trans. Wirel. Commun.*, vol. 15, no. 5, pp. 3718–3731, 2016, doi: 10.1109/TWC.2016.2526601.

[145] R. Di Pietro, L. V. Mancini, Y. W. Law, S. Etalle, and P. Havinga, "LKHW: A directed diffusion-based secure multicast scheme for wireless sensor networks," *Proc. Int. Conf. Parallel Process. Work.*, vol. 2003-Janua, no. 1, pp. 397–406, 2003, doi: 10.1109/ICPPW.2003.1240395.

[146] S. H. Seo, J. Won, S. Sultana, and E. Bertino, "Effective key management in dynamic wireless sensor networks," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 2, pp. 371–383, 2015, doi: 10.1109/TIFS.2014.2375555.

[147] M. Dong, K. Ota, L. T. Yang, A. Liu, and M. Guo, "LSCD: A Low-Storage Clone Detection Protocol for Cyber-Physical Systems," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 35, no. 5, pp. 712–723, 2016, doi: 10.1109/TCAD.2016.2539327.

[148] V. Pal, Yogita, G. Singh, and R. P. Yadav, "Cluster Head Selection Optimization Based on Genetic Algorithm to Prolong Lifetime of Wireless Sensor Networks," *Procedia Comput. Sci.*, vol. 57, no. Icrtc, pp. 1417–1423,

2015, doi: 10.1016/j.procs.2015.07.461.

[149] M. Ye, Y. Wang, C. Dai, and X. Wang, "A Hybrid Genetic Algorithm for the Minimum Exposure Path Problem of Wireless Sensor Networks Based on a Numerical Functional Extreme Model," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8644–8657, 2016, doi: 10.1109/TVT.2015.2508504.

[150] X. Li, L. Xu, H. Wang, J. Song, and S. X. Yang, "A differential evolution-based routing algorithm for environmental monitoring wireless sensor networks," *Sensors*, vol. 10, no. 6, pp. 5425–5442, 2010, doi: 10.3390/s100605425.

[151] P. Kuila and P. K. Jana, "A novel differential evolution based clustering algorithm for wireless sensor networks," *Appl. Soft Comput.*, vol. 25, pp. 414–425, 2014, doi: 10.1016/j.asoc.2014.08.064.

[152] Y. Lu, J. Chen, I. Comsa, P. Kuonen, and B. Hirsbrunner, "Construction of data aggregation tree for multi-objectives in wireless sensor networks through jump particle swarm optimization," *Procedia Comput. Sci.*, vol. 35, no. C, pp. 73–82, 2014, doi: 10.1016/j.procs.2014.08.086.

[153] M. Taherian, H. Karimi, A. M. Kashkooli, A. Esfahanimehr, T. Jafta, and M. Jafarabad, "The Design of an Optimal and Secure Routing Model in Wireless Sensor Networks by Using PSO Algorithm," *Procedia Comput. Sci.*, vol. 73, no. Awict, pp. 468–473, 2015, doi: 10.1016/j.procs.2015.12.028.

[154] X. Liu and D. He, "Ant colony optimization with greedy migration mechanism for node deployment in wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 39, no. 1, pp. 310–318, 2014, doi: 10.1016/j.jnca.2013.07.010.

[155] X. Sun, Y. Zhang, X. Ren, and K. Chen, "Optimization deployment of wireless sensor networks based on culture-ant colony algorithm," *Appl. Math. Comput.*, vol. 250, pp. 58–70, 2015, doi: 10.1016/j.amc.2014.10.091.

[156] D. Karaboga, S. Okdem, and C. Ozturk, "Cluster based wireless sensor network routing using artificial bee colony algorithm," *Wirel. Networks*, vol. 18, no. 7, pp. 847–860, 2012, doi: 10.1007/s11276-012-0438-z.

[157] J. Roselin and P. Latha, "Energy efficient coverage using artificial bee colony optimization in wireless sensor networks," *J. Sci. Ind. Res. (India).*, vol. 75, no. 1, pp. 19–27, 2016, doi: 10.1022/jsir/2016.22.4456.

[158] N. A. Al-Aboody and H. S. Al-Raweshidy, "Grey Wolf optimization-based energy-efficient routing protocol for heterogeneous wireless sensor networks," *2016 4th Int. Symp. Comput. Bus. Intell. ISCBI 2016*, no. 1, pp. 101–107, 2016, doi: 10.1109/ISCBI.2016.7743266.

[159] I. Grey, W. Optimizer, X. Zhao, H. Zhu, S. Aleksic, and Q. Gao, "Energy-Efficient Routing Protocol for Wireless Sensor Networks Based on Improved Grey Wolf Optimizer," *KSII Trans. Internet Inf. Syst.*, vol. 12, no. 6, pp.

2644–2657, 2018, doi: 10.3837/tiis.2018.06.011.

[160] M. Sharawi and E. Emary, "Impact of grey Wolf optimization on WSN cluster formation and lifetime expansion," *9th Int. Conf. Adv. Comput. Intell. ICACI 2017*, pp. 157–162, 2017, doi: 10.1109/ICACI.2017.7974501.

[161] G. P. Gupta and B. Saha, "Load balanced clustering scheme using hybrid metaheuristic technique for mobile sink based wireless sensor networks," *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2020, doi: 10.1007/s12652-020-01909-z.

[162] B. M. Sahoo, H. M. Pandey, and T. Amgoth, "GAPSO-H: A hybrid approach towards optimizing the cluster based routing in wireless sensor network," *Swarm Evol. Comput.*, vol. 60, no. January 2020, p. 100772, 2021, doi: 10.1016/j.swevo.2020.100772.

[163] Z. jun Teng, J. ling Lv, and L. wen Guo, "An improved hybrid grey wolf optimization algorithm," *Soft Comput.*, vol. 23, no. 15, pp. 6617–6631, 2019, doi: 10.1007/s00500-018-3310-y.

[164] Rekha and R. Gupta, "Cluster head election in wireless sensor network: A comprehensive study and future directions," *Int. J. Comput. Networks Appl.*, vol. 7, no. 6, pp. 178–192, 2020, doi: 10.22247/ijcna/2020/205319.

[165] E. Thenmozhi and S. Audithan, "Energy efficinet cluster head selection and data convening in wireless sensor networks," *Indian J. Sci. Technol.*, vol. 9, no. 15, 2016, doi: 10.17485/ijst/2016/v9i15/77749.

[166] D. Jia, H. Zhu, S. Zou, and P. Hu, "Dynamic cluster head selection method for wireless sensor network," *IEEE Sens. J.*, vol. 16, no. 8, pp. 2746–2754, 2016, doi: 10.1109/JSEN.2015.2512322.

[167] D. Agrawal and S. Pandey, "FUCA: Fuzzy-based unequal clustering algorithm to prolong the lifetime of wireless sensor networks," *Int. J. Commun. Syst.*, vol. 31, no. 2, pp. 1–18, 2018, doi: 10.1002/dac.3448.

[168] P. Neamatollahi and M. Naghibzadeh, "Distributed unequal clustering algorithm in large-scale wireless sensor networks using fuzzy logic," *J. Supercomput.*, vol. 74, no. 6, pp. 2329–2352, 2018, doi: 10.1007/s11227-018-2261-5.

[169] P. S. Mehra, M. N. Doja, and B. Alam, "Fuzzy based enhanced cluster head selection (FBECS) for WSN," *J. King Saud Univ. - Sci.*, vol. 32, no. 1, pp. 390–401, 2020, doi: 10.1016/j.jksus.2018.04.031.

[170] Y. Jeong, J.-H. Lee, and G.-S. Shin, "Based on Agile Methodology," *Adv. Commun. Technol. 2008. ICACT 2008. 10th Int. Conf.*, vol. 1, pp. 362–366, 2008, doi: 10.1109/ICACT.2008.4493779.

[171] A. Krishnakumar and V. Anuratha, "An energy-efficient cluster head selection of LEACH protocol for wireless sensor networks," *2017 Int. Conf.*

Nextgen Electron. Technol. Silicon to Software, ICNETS2 2017, pp. 57–61, 2017, doi: 10.1109/ICNETS2.2017.8067897.

[172] P. Azad and V. Sharma, "Clusterhead selection using multiple attribute decision making (MADM) approach in wireless sensor networks," *Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng. LNICST*, vol. 115, pp. 141–154, 2013, doi: 10.1007/978-3-642-37949-9_12.

[173] A. Tamizharasi, J. J. Selvathai, A. Kavi Priya, R. Maarlin, and M. Harinetha, "Energy aware heuristic approach for cluster head selection in wireless sensor networks," *Bull. Electr. Eng. Informatics*, vol. 6, no. 1, pp. 70–75, 2017, doi: 10.11591/eei.v6i1.598.

[174] R. A. Kovacs, B. Iancu, V. T. Dadarlat, E. Cebuc, and S. Buzura, "Extending K-cover genetic algorithm for efficient energy consumption in WSNs," *Proc. - RoEduNet IEEE Int. Conf.*, vol. 2019-Octob, 2019, doi: 10.1109/ROEDUNET.2019.8909581.

[175] M. Esmaeili, A. Branch, and S. Jamali, "Energy-Aware Clustering in the Internet of Things by Using the Genetic Algorithm," *J. Tele. Elect. Comp. Eng.,* vol. 12, no. 2, pp. 29–37, 2020, doi: 10.1009/*JTEC*/2020.29.5037.

[176] S. Tabibi and A. Ghaffari, "Energy-Efficient Routing Mechanism for Mobile Sink in Wireless Sensor Networks Using Particle Swarm Optimization Algorithm," *Wirel. Pers. Commun.*, vol. 104, no. 1, pp. 199–216, 2019, doi: 10.1007/s11277-018-6015-8.

[177] C. Wang, X. Liu, H. Hu, Y. Han, and M. Yao, "Energy-efficient and load-balanced clustering routing protocol for wireless sensor networks using a chaotic genetic algorithm," *IEEE Access*, vol. 8, pp. 158082–158096, 2020, doi: 10.1109/ACCESS.2020.3020158.

[178] T. Kaur and D. Kumar, "Particle Swarm Optimization-Based Unequal and Fault Tolerant Clustering Protocol for Wireless Sensor Networks," *IEEE Sens. J.*, vol. 18, no. 11, pp. 4614–4622, 2018, doi: 10.1109/JSEN.2018.2828099.

[179] B. M. Sahoo, T. Amgoth, and H. M. Pandey, "Particle swarm optimization based energy efficient clustering and sink mobility in heterogeneous wireless sensor network," *Ad Hoc Networks*, vol. 106, p. 102237, 2020, doi: 10.1016/j.adhoc.2020.102237.

[180] P. Kuila, S. K. Gupta, and P. K. Jana, "A novel evolutionary approach for load balanced clustering problem for wireless sensor networks," *Swarm Evol. Comput.*, vol. 12, pp. 48–56, 2013, doi: 10.1016/j.swevo.2013.04.002.

[181] A. Bari, A. Jaekel, and S. Bandyopadhyay, "Clustering strategies for improving the lifetime of two-tiered sensor networks," *Comput. Commun.*, vol. 31, no. 14, pp. 3451–3459, 2008, doi: 10.1016/j.comcom.2008.05.038.

[182] G. K. Ijemaru, K. L. M. Ang, and J. K. P. Seng, "Wireless power transfer and

energy harvesting in distributed sensor networks: Survey, opportunities, and challenges," *Int. J. Distrib. Sens. Networks*, vol. 18, no. 3, 2022, doi: 10.1177/15501477211067740.

[183] J. Amutha, S. Sharma, and J. Nagar, "WSN Strategies Based on Sensors, Deployment, Sensing Models, Coverage and Energy Efficiency: Review, Approaches and Open Issues," *Wirel. Pers. Commun.*, vol. 111, no. 2, pp. 1089–1115, 2020, doi: 10.1007/s11277-019-06903-z.

[184] P. Kumar and S. R. N. Reddy, "Wireless sensor networks: a review of motes, wireless technologies, routing algorithms and static deployment strategies for agriculture applications," *CSI Trans. ICT*, vol. 8, no. 3, pp. 331–345, 2020, doi: 10.1007/s40012-020-00289-1.

[185] S. Jabbar *et al.*, "Analysis of Factors Affecting Energy Aware Routing in Wireless Sensor Network," *Wirel. Commun. Mob. Comput.*, vol. 2018, 2018, doi: 10.1155/2018/9087269.

[186] P. S. Rathore, J. M. Chatterjee, A. Kumar, and R. Sujatha, "Energy-efficient cluster head selection through relay approach for WSN," *J. Supercomput.*, vol. 77, no. 7, pp. 7649–7675, 2021, doi: 10.1007/s11227-020-03593-4.

[187] V. Seedha Devi, T. Ravi, and S. B. Priya, "Cluster Based Data Aggregation Scheme for Latency and Packet Loss Reduction in WSN," *Comput. Commun.*, vol. 149, no. July 2019, pp. 36–43, 2020, doi: 10.1016/j.comcom.2019.10.003.

[188] R. Yarinezhad and S. N. Hashemi, "Solving the load balanced clustering and routing problems in WSNs with an fpt-approximation algorithm and a grid structure," *Pervasive Mob. Comput.*, vol. 58, p. 101033, 2019, doi: 10.1016/j.pmcj.2019.101033.

[189] D. W. Sambo, B. O. Yenke, A. Förster, and P. Dayang, "Optimized clustering algorithms for large wireless sensor networks: A review," *Sensors (Switzerland)*, vol. 19, no. 2, 2019, doi: 10.3390/s19020322.

[190] S. Arora and S. Singh, "Butterfly optimization algorithm: a novel approach for global optimization," *Soft Comput.*, vol. 23, no. 3, pp. 715–734, 2019, doi: 10.1007/s00500-018-3102-4.

[191] A. S. Assiri, "On the performance improvement of butterfly optimization approaches for global optimization and Feature Selection," *PLoS One*, vol. 16, no. 1 January, 2021, doi: 10.1371/journal.pone.0242612.

[192] P. Sivakumar and M. Radhika, "Performance Analysis of LEACH-GA over LEACH and LEACH-C in WSN," *Procedia Comput. Sci.*, vol. 125, pp. 248–256, 2018, doi: 10.1016/j.procs.2017.12.034.

[193] H. H., "Performance comparison of LEACH, SEP and Z-SEP Protocols in WSN," *Int. J. Comput. Appl.*, vol. 180, no. 30, pp. 41–46, 2018, doi: 10.5120/ijca2018916780.

[194] P. Maheshwari, A. K. Sharma, and K. Verma, “Energy efficient cluster based routing protocol for WSN using butterfly optimization algorithm and ant colony optimization,” *Ad Hoc Networks*, vol. 110, no. May 2020, p. 102317, 2021, doi: 10.1016/j.adhoc.2020.102317.

[195] T. Bhatia, S. Kansal, S. Goel, and A. K. Verma, “A genetic algorithm based distance-aware routing protocol for wireless sensor networks,” *Comput. Electr. Eng.*, vol. 56, pp. 441–455, 2016, doi: 10.1016/j.compeleceng.2016.09.016.

[196] P. Lalwani, I. Ganguli, and H. Banka, “FARW: Firefly algorithm for Routing in wireless sensor networks,” *2016 3rd Int. Conf. Recent Adv. Inf. Technol. RAIT 2016*, pp. 248–252, 2016, doi: 10.1109/RAIT.2016.7507910.

[197] N. Jain and M. Mannan, “Comparative Performance Analysis of Teen Sep Leach Eammh and Pegasis Routing Protocols,” *Int. Res. J. Engin. Tech.,* pp. 983–987, 2016, doi:10.1109/ IRJET.2016.7543203.

[198] M. S. Manshahia, M. Dave, and S. B. Singh, “Firefly algorithm based clustering technique for Wireless Sensor Networks,” *Proc. 2016 IEEE Int. Conf. Wirel. Commun. Signal Process. Networking, WiSPNET 2016*, pp. 1273–1276, 2016, doi: 10.1109/WiSPNET.2016.7566341.

[199] S. Vancin and E. Erdem, “Performance analysis of the energy efficient clustering models in wireless sensor networks,” *ICECS 2017 - 24th IEEE Int. Conf. Electron. Circuits Syst.*, vol. 2018-Janua, pp. 247–251, 2018, doi: 10.1109/ICECS.2017.8292040.

[200] F. Fanian and M. Kuchaki Rafsanjani, “Cluster-based routing protocols in wireless sensor networks: A survey based on methodology,” *J. Netw. Comput. Appl.*, vol. 142, no. February, pp. 111–142, 2019, doi: 10.1016/j.jnca.2019.04.021.

[201] R. Sharma, V. Vashisht, and U. Singh, “Nature inspired algorithms for energy efficient clustering in wireless sensor networks,” *Proc. 9th Int. Conf. Cloud Comput. Data Sci. Eng. Conflu. 2019*, pp. 365–370, 2019, doi: 10.1109/CONFLUENCE.2019.8776618.

[202] Z. Sun, M. Wei, Z. Zhang, and G. Qu, “Secure Routing Protocol based on Multi-objective Ant-colony-optimization for wireless sensor networks,” *Appl. Soft Comput. J.*, vol. 77, pp. 366–375, 2019, doi: 10.1016/j.asoc.2019.01.034.

[203] S. Sankar, S. Ramasubbareddy, F. Chen, and A. H. Gandomi, “Energy-Efficient Cluster-based Routing Protocol in Internet of Things Using Swarm Intelligence,” *2020 IEEE Symp. Ser. Comput. Intell. SSCI 2020*, pp. 219–224, 2020, doi: 10.1109/SSCI47803.2020.9308609.

[204] P. Nandhini and A. Suresh, “Energy Efficient Cluster Based Routing Protocol Using Charged System Harmony Search Algorithm in WSN,” *Wirel. Pers. Commun.*, vol. 121, no. 3, pp. 1457–1470, 2021, doi: 10.1007/s11277-021-08679-7.

[205] R. Xie and X. Jia, "Transmission-efficient clustering method for wireless sensor networks using compressive sensing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 3, pp. 806–815, 2014, doi: 10.1109/TPDS.2013.90.

[206] D. Fotue, F. Melakessou, H. Labiod, and T. Engel, "Mini-sink mobility with diversity-based routing in wireless sensor networks," *PE-WASUN'11 - Proc. 8th ACM Symp. Perform. Eval. Wirel. Ad Hoc, Sensor, Ubiquitous Networks*, pp. 9–16, 2011, doi: 10.1145/2069063.2069066.

[207] A. Goyal, S. Mudgal, and S. Kumar, "A Review on Energy-Efficient Mechanisms for Cluster-Head Selection in WSNs for IoT Application," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1099, no. 1, p. 012010, 2021, doi: 10.1088/1757-899x/1099/1/012010.

[208] M. N. U. Islam, A. Fahmin, M. S. Hossain, and M. Atiquzzaman, "Denial-of-Service Attacks on Wireless Sensor Network and Defense Techniques," *Wirel. Pers. Commun.*, vol. 116, no. 3, pp. 1993–2021, 2021, doi: 10.1007/s11277-020-07776-3.

[209] G. Liu, H. Zhao, F. Fan, G. Liu, Q. Xu, and S. Nazir, "An Enhanced Intrusion Detection Model Based on Improved kNN in WSNs," *Sensors*, vol. 22, no. 4, pp. 1–18, 2022, doi: 10.3390/s22041407.

[210] A. Khanna and S. Kaur, "Internet of Things (IoT), Applications and Challenges: A Comprehensive Review," *Wirel. Pers. Commun.*, vol. 114, no. 2. Springer US, 2020, doi: 10.1007/s11277-020-07446-4 .

[211] T. D. Diwan, S. Choubey, and H. S. Hota, "A Detailed Analysis on NSL-KDD Dataset using various Machine Learning Techniques for Intrusion Detection," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 11, pp. 2954–2968, 2021, doi: 10.1108/ TURCOMAT/2021.19291.

[212] K. Siddique, Z. Akhtar, F. Aslam Khan, and Y. Kim, "KDD Cup 99 Data Sets: A Perspective on the Role of Data Sets in Network Intrusion Detection Research," *Computer (Long. Beach. Calif).*, vol. 52, no. 2, pp. 41–51, 2019, doi: 10.1109/MC.2018.2888764.

[213] L. You, Y. Li, Y. Wang, J. Zhang, and Y. Yang, "A deep learning-based RNNs model for automatic security audit of short messages," *2016 16th Int. Symp. Commun. Inf. Technol. Isc. 2016*, pp. 225–229, 2016, doi: 10.1109/ISCIT.2016.7751626.

[214] B. Dong and X. Wang, "Comparison deep learning method to traditional methods using for network intrusion detection," *Proc. 2016 8th IEEE Int. Conf. Commun. Softw. Networks, ICCSN 2016*, pp. 581–585, 2016, doi: 10.1109/ICCSN.2016.7586590.

[215] M. Attia, S. M. Senouci, H. Sedjelmaci, E. H. Aglzim, and D. Chrenko, "An efficient Intrusion Detection System against cyber-physical attacks in the smart grid," *Comput. Electr. Eng.*, vol. 68, no. September 2017, pp. 499–512, 2018, doi: 10.1016/j.compeleceng.2018.05.006.

[216] S. Yılmaz Gündüz and M. N. ÇETER, "Feature Selection and Comparison of Classification Algorithms for Intrusion Detection," *ANADOLU Univ. J. Sci. Technol. A - Appl. Sci. Eng.*, vol. 19, no. 1, pp. 206–218, 2018, doi: 10.18038/aubtda.356705.

[217] A. A. Aburomman and M. B. I. Reaz, "A survey of intrusion detection systems based on ensemble and hybrid classifiers," *Comput. Secur.*, vol. 65, pp. 135–152, 2017, doi: 10.1016/j.cose.2016.11.004.

[218] A. Mushtaq and H. Kanth, "Hybrid Wireless Network Approach for QoS," *Academia.Edu*, vol. 2, no. 5, pp. 1188–1190, 2015, [Online]. Available: http://www.ijritcc.org.

[219] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process.*, vol. 115, pp. 213–237, 2019, doi: 10.1016/j.ymssp.2018.05.050.

[220] H. Ao, "Using Machine Learning Models to Detect Different Intrusion on NSL-KDD," *2021 IEEE Int. Conf. Comput. Sci. Artif. Intell. Electron. Eng. CSAIEE 2021*, pp. 166–177, 2021, doi: 10.1109/CSAIEE54046.2021.9543241.

[221] S. Pande, A. Khamparia, and D. Gupta, "Feature selection and comparison of classification algorithms for wireless sensor networks," *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2021, doi: 10.1007/s12652-021-03411-6.

[222] C. Ozgur, T. Colliau, G. Rogers, and Z. Hughes, "MatLab vs. Python vs. R," *J. Data Sci.*, vol. 15, no. 3, pp. 355–372, 2021, doi: 10.6339/jds.201707_15(3).0001.

[223] D. H. Deshmukh, T. Ghorpade, and P. Padiya, "Improving classification using preprocessing and machine learning algorithms on NSL-KDD dataset," *Proc. - 2015 Int. Conf. Commun. Inf. Comput. Technol. ICCICT 2015*, pp. 13–18, 2015, doi: 10.1109/ICCICT.2015.7045674.

[224] B. Ingre and A. Yadav, "Performance analysis of NSL-KDD dataset using ANN," *Int. Conf. Signal Process. Commun. Eng. Syst. - Proc. SPACES 2015, Assoc. with IEEE*, pp. 92–96, 2015, doi: 10.1109/SPACES.2015.7058223.

[225] H. Benaddi, K. Ibrahimi, and A. Benslimane, "Improving the Intrusion Detection System for NSL-KDD Dataset based on PCA-Fuzzy Clustering-KNN," *Proc. - 2018 Int. Conf. Wirel. Networks Mob. Commun. WINCOM 2018*, 2019, doi: 10.1109/WINCOM.2018.8629718.