



Proceeding Paper A Review of Machine Learning-Based Routing Protocols for Wireless Sensor Network Lifetime ⁺

Abhay R. Gaidhani^{1,*} and Amol D. Potgantwar²

- ¹ Department of Computer Science and Engineering, MET Institute of Engineering, Affiliated to Savitribai Phule Pune University (SPPU), Nashik 422002, Maharashtra, India
- ² Department of Computer Engineering, Sandip Institute of Technology and Research Centre, Affiliated to Savitribai Phule Pune University (SPPU), Nashik 422002, Maharashtra, India; amol.potgantwar@sitrc.org
- * Correspondence: abhay.gaidhani2@gmail.com
- Presented at the International Conference on Recent Advances in Science and Engineering, Dubai, United Arab Emirates, 4–5 October 2023.

Abstract: Wireless sensor networks (WSNs) grapple with a challenging and pivotal issue: how to maximize the network's lifespan. To improve the quality of service (QoS) and extend the life of the network, there has been a lot of effort made in this area in recent years. The sensor nodes of a WSN are autonomous, dispersed devices that gather and direct data to a central hub, or "Base Station", using wireless connections without any central coordinator. These networks have less processing power, memory capacity, power supply, and so on, which limits their range and battery life. Numerous studies preceding our research have proposed a myriad of strategies to enhance network longevity. In a WSN, information is relayed from one node to the next until it reaches the base station. Most nodes may be reliably expected to operate for the duration of their batteries. These strategies encompass reducing energy consumption, minimizing latency, load balancing, clustering, efficient data aggregation, and curtailing data transmission delays. WSNs may alter dynamically as a result of internal or external circumstances, necessitating a depreciating dispensable redesign of the network. Because the networks in classic WSN techniques are expressly programmed, it is difficult for them to respond dynamically. Machine learning (ML) approaches can be used to respond appropriately in order to overcome such circumstances. Machine learning is the process of acting without human involvement or reprogramming in order to learn from experiences. This paper presents a review of different ML-based algorithms for WSNs together with their benefits, limitations, and parameters that affect network longevity.

Keywords: machine learning; wireless sensor network; routing; neural network; fuzzy logic; reinforcement learning

1. Introduction

To improve the quality of service (QoS) and extend the life of the network, there has been a lot of effort made in this area in recent years. The sensor nodes of a WSN are autonomous, dispersed devices that gather and direct data to a central hub, or "Base Station", using wireless connections without any central coordinator. These networks have less processing power, memory capacity, power supply, and so on, which limits their range and battery life. In a WSN, information is relayed from one node to the next until it reaches the base station. Most nodes may be reliably expected to operate for the duration of their batteries. Investing in costly and time-consuming maintenance on sensor nodes in an area is not an option. It is possible to connect many devices directly; however, doing so may drastically shorten the network's lifespan. Improving a WSN's energy efficiency or prolonging its lifespan is a challenging topic that can be solved by a variety of innovations at various levels of the network's protocol stack. Effective congestion detection algorithms, energy-efficient routing, and power-efficient media access control all



Citation: Gaidhani, A.R.; Potgantwar, A.D. A Review of Machine Learning-Based Routing Protocols for Wireless Sensor Network Lifetime. *Eng. Proc.* **2023**, *59*, 231. https:// doi.org/10.3390/engproc2023059231

Academic Editors: Nithesh Naik, Rajiv Selvam, Pavan Hiremath, Suhas Kowshik CS and Ritesh Ramakrishna Bhat

Published: 20 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). play a role. Due to the significance of node placement in WSNs, a few research publications have developed mathematical formulations for the best possible node distribution schemes. While some studies focused on using dispatch, a node's sole purpose is to relay packets rather than do recognizing, and others emphasized the need to strategically place relay nodes at various distances from the master station. An energy-effective WSN design considers factors such as cluster-head selection, cluster size, and the frequencies with which cluster member characteristics are updated. Although some employ on/off switches, others still aggregate data to find patterns that can lead to energy savings. It is possible that the time and energy put into finding a solution to this problem can be reduced by using the techniques described in [1]. Similarly, "hierarchical or cluster-based routing" would be a particularly useful methodology for more scale WSNs to address the challenges of scalability, effective communication, and fault acceptance. Clusters are subnetworks organized in a hierarchical structure.

2. The Characteristics of WSNs

Computing abilities: Due to its small size, high price, and high battery consumption, the sensor has very limited computational capabilities. In terms of battery energy, sensor nodes frequently become invalid and out of control due to battery fatigue. So, the conservation protocols and algorithms need to be thought about ahead of time.

Communication capabilities: Due of their restricted communication capabilities and range, WSN communication devices have limited utility. The typical communication range is less than a meter, and environmental factors can have an impact on the sensor.

Dynamic Power Management (DPM): Energy consumption at the sensor node can be managed and preserved with DPM. Batteries power the sensor nodes, and how much juice they need varies with the task at hand (data collection, processing, or transmission). Because it is difficult to charge or replace the batteries, the sensor network cannot be used for a longer period of time.

Multi-hop communications: transmissions in a WSN occur over multiple hops in a radio frequency chain, allowing for intercommunication between nodes.

Application relevance: Data aggregation from WSNs is used to analyze traffic patterns and other factors. The sensor network deviates from the standard network in that it does not rely on a central server to collect data but rather on the apps running on it and the final working environment. The physical signals and routing protocols used by the various sensor network applications are not shared.

Reliability: the packet dimension, node frequency, protocol architecture, hop-by-hop connectivity, endways communication, buffer size, packet loss ratio, packet distribution proportion, bandwidth, and loss retrieval are just some of the aspects that affect WSN consistency [2].

Reliability semantics: Complete packet delivery is what packet reliability ensures. To fully define the reliability semantics, it is necessary to consider not only event reliability but also end-to-end consistency, upstream consistency, and downstream reliability.

General problems in WSNs: In terms of node placement, initial and foremost is the difficulty of getting the nodes set up, which is complicated by factors like cost and locational restrictions [3].

Energy consumption without accuracy: establishing a connection between nodes and base stations (BSs) requires a certain number of nodes, the consumption of which might influence the network's lifespan.

Coverage: the sensor's narrow field of view means it can only monitor a subset of the network's physical space.

Node dynamics: Nodes in networks are presumed stable. In other circumstances, however, it becomes critical to fortify the adaptability of base stations or CHSs. Since node membership tends to be fluid and the cluster is subject to change over time, studying node dynamics presents a formidable challenge.

Overlay cluster: As was previously noted, the network administrator or the sensors themselves can decide on the CH. A person who is currently part of one cluster is likely to switch over to another CH if they make their decision at a later date. To address design concerns, it cuts across functional silos. In order to avoid inequity, starvation, or stagnation because of resource rivalry, it is crucial to put in place the mechanisms that will allow for the detection of overlapping groups and the coordination between them.

Load balancing: In sensor networks, this is a serious issue because CHs are typically created after the initial deployment of sensors. The head will overload the specific CH responsible for failure if the delivery of member sensor nodes is not balanced over all available CHs. So, it is important to form groups of similar size in such circumstances to ensure a level playing field.

3. Routing on Wireless Sensor Networks

By controlling the power usage of individual nodes, routing can help a network last longer under specific situations. WSNs employ several protocols whose primary focus is on extending the operational lifetime of the network. What determines this is the network's architecture and the routes it takes. In this case, we used a mixed method with the first ever location-based routing. The other option is to route groups together. The structure of the internet lends credence to this categorization.

3.1. Routing Approach

It was discovered that such a strategy falls into two distinct methodological camps, i.e., the following:

Classical-based routing: By selecting nodes at random using a timed function, traditional routing provides an unstable flow of traffic between the many nodes in the core cluster, which in turn causes problems with scalability, load balancing, connection, and coverage.

Optimized-based routing: This demonstrates a critical function in prolonging the WSN network's usefulness. The key problems with this system are its intolerance, poor energy efficiency, poor connectivity, rigidity, and inability to scale.

3.2. Clustering Characteristics

Clustering characteristics entail how certain aspects of the clustering procedure can affect the effectiveness of hierarchical routing. Cluster qualities have two major issues: cluster characteristics and sensor capacities [4–6].

3.3. Cluster Purposes

Improving the lifetime of the network: the success of an SN deployment is highly dependent on the network's longevity; as a result, it is important to study routing algorithms that minimize energy waste while maximizing throughput.

Fault tolerance: Many methods are developed to deal with the problem of a node failing. Proxy CHs are used or CH rotation is used (when both nodes fail) to deal with this problem. When developing the clustering algorithm, fault tolerance was one of the primary considerations.

Fill handling: When outlining the goals of a clustering protocol's design, fill handling is an important one. Overloading the CHs will quickly drain their energy reserves and is therefore strictly prohibited. Thus, it is crucial that each cluster's nodes be evenly distributed as CHs carry out the work of data aggregation or signal processing tasks.

3.4. Cluster-Head Choice Norms

The following criteria are considered when picking a CH:

While the choice of CH has a significant impact on the lifespan of WSNs, it is important to take the network's topology into account when making this decision.

Residual energy (RE): Its power is crucial for the CH selection procedure. It appears that this power is all that is left in the sensor once it has completed its communication. This RE reduces the energy expenditure of the node while the CH and BS are exchanging information.

3.5. Cluster Properties

Cluster size: On total available nodes and the network transmission range, the lifetime of a WSN can be affected by its cluster size. Increasing the count of nodes in a cluster may lead to a decrease in total available node energy. More iterations may cause node energy depletion if there are a more nodes in the cluster.

Cluster density: Energy consumption in CHs is proportional to the total number of cluster members. The cluster's density can change depending on whether it is a static or dynamic cluster.

3.6. Detailed Procedure of Clustering Routing Protocols

- (1) *Finalizing the number of cluster heads*: after establishing the clustering number K with the scientific model and formula, the clustering probability, P = K/N, can be calculated.
- (2) Initial clustering: To begin, the network must be initialized and initially clustered with LEACH, with the aim of identifying the temporary clusters and clustering of the head. At the same time, the temporary CH gathers data on the energy and location of each cluster node.

3.6.1. Fuzzy Logic

The four main components of a fuzzy logic system (FLS) are the fuzzifier, the inferencing engine, the rules-based decision-making, and the fuzzifier. Entering the data and utilizing the existing purposes to recognize the grade of each applicable fuzzy set are the first step. The term "fuzzy inference" refers to the application of fuzzy logic to the task of drawing conclusions about an input in terms of an expected output. The depiction then yields recommendations for model identification or decision-making. When a fuzzy inference system generates an imprecise result, this is referred to as degradation.

FCH: Fuzzy-logical CH selection (FCH) is performed [7]. It utilizes energy and node cognizance and focuses on the period of calculation to make a CH selection using an integrated procedure, FL.

CHEF: In the paper [8], the authors propose a fuzzy-logic-based CH selection mechanism (CHEF). However, CHEF is used to book other nodes for BS and CH choices, as opposed to LEACH's "hot spare" approach.

LEACH-FL: As stated in [9], LEACH-FL is an improved type of the LEACH algorithm. It employs a device like that of the LCH, albeit with parameters tailored to things like battery life, proximity to the sink, and network density.

ICT2TSK: Type 2 "Takagi-Sugeno-Kang (TSK) as the FL system" [10] is used in "ICT2TSK", a sophisticated clustering algorithm. When compared to the TSK FL Type 1 system, it is used to pick CHs and more effectively accomplish the governing of uncertainties. By instituting a uniform zone of competition around each CH, it can gauge the load on the network and make informed decisions about how best to allocate power.

SEP-FL: In a heterogeneous WSN, SEP-FL is a FL method that improves CH selectivity [11]. By adjusting the probabilities of each node's residual energy, this refined form of SEP makes use of CH selectivity.

EAUCF: In [12], the author suggests using a fuzzy algorithm called Energy-Aware in an Unequal Cluster (EAUCF).

DFLCR: To address the problems of network period and energy consumed in WSNs, the writer suggests a dispersed fuzzy-logic-based clustering routing (DFLCR) [13] protocol. In contrast to other protocols, the suggested DFLCR protocol considers the significance of local distance and reserved residual energy. Because the DFLCR protocol can regulate the cluster size, the best node can be chosen as the CH., Fuzzy logic (FL) has been used to address a wide range of issues plaguing WSNs.

SIF: this is a swarm-intelligence protocol based on FL (SIF) routing and RE [14], sink distance, and distance to center to choose appropriate CHs.

FBUC: An improvement on the EAUCF [15] is the fuzzy-based asymmetric UC (FBUC) method of grouping. According to FNB, FBUC outclasses the other two systems across all stages in terms of energy consumption and network lifetime.

EEDCF: When it comes to consuming less energy during clustering-based routing operations and cluster-head identification, the EEDCF [16] protocol, also known as "Energy-Efficient Distributed Clustering Based on Fuzzy" [16], is a differential algorithm in WSN.

3.6.2. Genetic Algorithms (GAs)

A more efficient approach would involve using a GA to calculate the routing topology of the WSN again in the present while taking each node's energy level and distance from other nodes into account. Fitness tests are run on the new structures to select the fittest persons in order to enhance the GA models. This optimization technique benefits from the information from random searches and routes. GAs have demonstrated their ability to adapt to changing conditions, making them deployable. The six basic elements of the GA global scheme are the preliminary set of outlines, chromosomal ciphering task, gene operator, assessment function, choice method, and final set of configurations.

The "genetic algorithm to extend the lifespan of sensor networks at two levels (GABEEC)" of Wazed et al. is proposed in [17]. The GABEEC algorithm is proposed by Bayrakli et al. in [18]. The protocol has two phases: the first is the configuration phase, and the second is the steady-state phase. Based on energy entropy, Hamidouche et al. [19] propose a genetic process for multipath routing in WSNs (GAEMW). In order to boost lifetime efficiency and WSN steadiness, the authors of [20], Abo-Zahhad et al., disseminated an additional adaptive energy-efficacy clustering hierarchy protocol (GAEEP) constructed in a genetic process. The protocol's goal is to rise the valuable lifetime of WSNs by determining the best placement for CHs and the minimum possible number of them to use, all while decreasing the energy requirement of separate sensor nodes. The goal of the research presented by Hussain et al. in [21] was to extend the life of networks and use less energy. Here, we use the constraints of "BS distance, data transfer, cluster distance, and energy level" [21]. We find that LEACH, HCR-1, and HCR-2 all perform at or near the same level. According to the outcomes, optimizing the fraction of the network that is actively processing data is more effective than using more algorithms.

In order to attain optimum thresholding for cluster construction, the authors Liu et al. suggest a GA-based adaptive clustering algorithm. LEACH-GA is a GA-based enhancement of the original LEACH [22]. Moreover, the LEACH-GA technique outperforms DT, LEACH, and MTE in terms of network lifetime, and its use of the finest possibility accepts optimum energy-efficient clustering. This is in comparison to the LEACH configuration and static state. The suggested algorithm calculates the best clustering threshold. The findings demonstrate that, in comparison to leech-GA leeches, the network lifetime can be better quality by growing the number of live nodes per cycle. Complex systems known as Artificial Neural Networks (ANNs) use components referred to as neurons or nerve cells to carry out parallel or distributed processing. Synapses, which are weighted connections, are used for communication between neurons. ANNs are computational processes that can generate principal multifaceted mapping involving different parameters by means of a supervised learning method in order to finish the categorization of data using an unsupervised manner.

4. Estimation of the Enhanced Clustering Methods

Many clustering algorithms that take machine learning into account in terms of key parameters like data delivery speediness, nonstop power consumption over the time of a network, and algorithm scalability are compared here. The uniformity of nodes, whether or not data collection techniques are integrated, and whether or not the system is centralized all play a role. The rate at which information is transferred from sending nodes to a base station (BS) or other receiving nodes is a critical metric, as it provides an overarching view of the transmission's reliability and the amount of data lost. When evaluating clustering algorithms, one crucial metric is the energy-consumption parameter. Reduction in node energy impacts the lifetime of the sensor network because more communication or transactions at a node can use up energy and decrease the lifespan of the network. As an outcome, we have focused our research on WSNs on a grander scale, debating the scalability of existing algorithms and suggesting that networks with more than 500 nodes receive special attention.

There are two main methods for implementing clustering algorithms: a centralized method and an extensive method. However, in the centralized method, the BS is an integral part of the decision-making process. In this context, homogeneity in a distributed system refers to a set of nodes in the network that all share the same characteristics, and they require that a BS broadcast this information to its plotting nodes and compute it as part of the ML adaptation to the network. In such a network, the CH is chosen as a clustering algorithm for nodes that have distinctive properties. The sensor network's radio model provides a rough estimate of the node's power consumption in the field. As such, the clustering method takes into account this radio model as part of the network and determines the energy output of each node during transmission. To measure how well the provided methods cope with unexpected situations, we define multipath metrics. Errors are tolerated better in multipath implementation techniques, unless the direct route to the BS is cut off and only an alternate route leads to the BS. This configuration allows the algorithm to continue operating normally in the event that a single node in the network fails.

5. Machine Learning Adaptability for WSNs

Many issues plaguing WSNs can be easily overcome and performance can be greatly improved by utilizing machine learning. Figure 1 shows some ways in which ML is used to extend the lifespan of WSNs include the following:

- Anticipating the volume of energy that can be collected during a given time slot, which
 is necessary to increase the longevity of WSNs;
- ML-based localization that actually works;
- Using ML to identify the faulty links;
- Using ML to decrease the dimensionality of data at the level of the sensor or CH.



Figure 1. Classification of ML techniques.

5.1. Supervised Learning

The supervised learning technique is used to fix several WSN issues, including localization [23], coverage difficulties [24], data collection [25], event discovery [26], routing [27], and target tracking [28]. The following are some supervised learning techniques:

Regression: As a supervised learning technique, regression makes predictions about a given value (Y) based on a predetermined collection of attributes (X). The regression model uses continuous or quantitative variables. Regression is a relatively basic machine learning technique that produces precise predictions with few errors.

Support vector machines: For optimization problems with undefined constraints or complex curves, SVMs provide an alternative solution. When packets cannot pass through the network near the sink nodes, as happens in a WSN, the sink is congested. In machine learning, the "support vector machine (SVM)" is a valuable classification process. Localization issues at the border and coverage gaps can be effectively addressed with SVM [29].

Decision trees: "Decision tree" refers to a learning algorithm that uses tree-like graphs to practice and evaluate separate functions. Using a decision tree to pick the best node is one of the most important steps in increasing the network lifetime of WSNs [30]. This is accomplished through the CH selection process. In addition, decision trees can be used as a strategy and a technique for localizing important features of linkage reliability, such as loss rate, normal retrieval time, and normal failure time.

Random forest: A collection of trees and a classification provided by each tree in the forest make up the random forest (RF) algorithm, which is a supervised machine learning technique. The random forest classifier is created first, and then the outcomes are predicted, which is how the RF algorithm operates [30]. For larger datasets and heterogeneous data, RF performs well. The missing values are correctly predicted using this method. A significant number of decision trees can be produced by isolating variables at each tree node and randomly choosing a subset of training samples. Because of the high quality of training data and overly robust decision trees, the RF classifier's sensitivity level is lower when compared to other streamlined machine learning classifiers.

Artificial Neural Networks: As the most popular learning algorithms, NNs are constructed from learning units called decision units. Perceptual and radial basis functions are common names for cascade circuits. Chains of decision units acting in concert can reveal both simple and intricate patterns in data. However, the computationally intensive learning practice requires multiple cascade circuits [31].

Deep learning: Deep learning is a subclass of ANNs and is a supervised machine learning technique used for categorization. The data-learning representation techniques with multi-layer representations (between the input and output layers) are known as deep learning approaches. In order to arrive at the optimal solution, it is composed of basic nonlinear modules that translate the representation from a lower layer to a higher layer. It draws inspiration from the information processing and communication patterns found in human nervous systems [32]. The abilities to extract high-level characteristics from data, work with or without labels, and be trained to accomplish numerous goals are the main advantages of deep learning.

K-nearest neighbor (*k*-*NN*): " k-NN is a supervised learning algorithm" in which the labels of nearby data models are used to classify sample test data. Nearest-average calculations are used to replace missing or unknowable test samples. The *k*-*NN* fault detection algorithm relies heavily on this method [31,32].

Bayesian learners: Without access to the mobility information contained in the WSN sensor node, Bayesian language learners [33] are able to make reliable estimates of sensor node movement (speed). They can create analytical mobility prediction models and routing-based mobility prediction methods using these mobility estimation techniques.

5.2. Unsupervised Learning

In this, there is no predetermined training model that is used to acquire knowledge; rather, it learns as it goes along. Moreover, no designations are offered. WSN utilizes these algorithms for data clustering and aggregation [34]. "K-means" plays a key role in "clustering and component analysis".

Principal Component Analysis: To improve the scalability of the network and reduce the overall power consumption, data aggregation plays a significant role in WSNs. The "Principal Component Analysis (PCA)" [35] is an appropriate technique for pre-processing data from sensor networks, i.e., for performing dimensionality decreases. To minimize energy consumption and maximize network lifetime, the PCA-guided routing algorithm explores the principal component analysis (PCA) technique to determine the shortest paths between nodes and cluster centers in a WSN.

K-means clustering: When it comes to partition-based algorithms for clustering data, "K-means clustering algorithmic [36]" is one of the most widely used models. By prioritizing the CHs with the highest energy, the K-means algorithm ensures that the clusters are well-balanced and have a low overall energy footprint [37–39].

Hierarchical clustering: Similar objects are grouped together into clusters using the hierarchical clustering technique, which determines the top-down or bottom-up order of each cluster. Divisive clustering, another name for top-down hierarchical clustering, involves splitting a huge single division recursively until there is one cluster for each observation. Agglomerative clustering, also known as bottom-up hierarchical clustering, allocates each observation to its cluster based on density functions. The hierarchical clustering approach is simple to use and does not require any prior knowledge of the number of groups.

Fuzzy-c-means clustering: Bezdek created fuzzy-c-means (FCM) clustering, also known as soft clustering, in 1981. It makes use of fuzzy set theory to assign an observation to one or more clusters. Using similarity measurements like intensity, distance, or connectedness, this technique finds clusters. The algorithms may be taken into consideration for one or more similarity measures, depending on the applications or datasets. To determine the ideal cluster centers, the algorithm runs over the clusters iteratively. Regarding the overlapped datasets, FCM yields the best grouping when compared to K-means.

5.3. Reinforcement Learning

In Table 1 many WSN initiatives, including collaborative communications, routing, and flow control, use "Reinforcement Learning (RL)" to guarantee that sensors and detecting nodes can observe optimum operations and execution networks in a suitable operating environment. There has been further development of the app's features. The term "trial and error" refers to the way that knowledge is gained via experience rather than through theoretical models. This model is notable for how easily it can be modified and put into practice. Processing and memory requirements are minimal, yet the potential applications are high. As such, RL for WSNs is concerned with a wide range of issues, including clustering protocols, medium-access protocols, and routing protocols; additionally, hyper-parameter optimization is addressed within the context of RL, with applications including architectural network design and application-level collaboration between nodes. To improve the effectiveness of standard learning algorithms for WSN routing issues, the Q-learning [40] approach is useful [41].

ML. Technique	Mobility	Environment	Complexity	Topology	Remarks
Regression	Static	Distributed/ Centralized	Moderate	Hybrid	Improved network lifetime
k-NN	Static	Distributed	Moderate	Hybrid	Reliable data transmission
Decision tree	Static	Distributed	Moderate	Tree	Improved network lifetime
ANN	Static Static	Centralized Distributed	High High	Hybrid Tree	Improved network lifetime
Decision tree	Static	Distributed	High	Tree	Reliable data transmission
Bayesian	Static or mobile	Distributed/ Centralized	Moderate	Hybrid	Improved network lifetime Reliable data transmission Improved time complexity Improved accuracy
k-means	Static	Distributed	Moderate	Tree	Efficient redundancy elimination
Hierarchical Clustering	Static	Distributed or Centralized	Moderate	Hierarchical	Reduce unnecessary transmission
РСА	Static	Distributed	Low Moderate High	Hybrid	Improved network lifetime Efficient drift findings Reduce unnecessary transmission
SVD	Static	Centralized	High	Hybrid	Reduce unnecessary transmission
Genetic Classifier	Static or mobile	Distributed	Moderate	Star	Improved network lifetime

Table 1. Comparison of ML-based techniques.

6. Limitations

Despite its many advantages, adopting ML in WSNs has certain limitations, including the following:

- Because machine learning (ML) algorithms need to learn from prior data, they do not yield correct predictions right away. The quantity of previous data affects the system's performance. Large data sizes also result in excessive energy consumption during data processing. Stated differently, a trade-off exists between the high computational complexity of the ML algorithm and the energy limits of WSNs. The ML algorithms must operate centrally in order to overcome this trade-off.
- It is a laborious process to validate the predictions made by the machine learning algorithm in a real-time setting.
- Finding a specific machine learning method to solve a problem in a wireless sensor network (WSN) can be hard at times.

7. Open Issues

There are still a number of unresolved issues that need more WSN research. Here, we have enumerated a few open research problems related to WSNs that may be resolved with machine learning technique:

Localization: Creating effective route planning for the beacon nodes within the framework of mobile sensor nodes is crucial. As far as we are aware, mobile anchor nodes do not have access to any specific established path-planning approach. Moreover, ML has the potential to offer effective route-planning methods for all sensor network anchor nodes, leading to improved localization accuracy and reduced power consumption. While the majority of current localization algorithms only consider two-dimensional space, the majority of real-time applications deploy sensor nodes in three-dimensional space. As a result, both stationary and mobile WSNs require the creation of methods for three-dimensional space localization.

Coverage and connectivity: Although many new approaches to coverage and connectivity have been suggested, many unanswered questions remain. Another difficult problem is determining the ideal location for sensor nodes and predicting the minimal number of nodes needed to cover a specific area. The majority of sensor node deployments in real-time WSN systems are haphazard. A coverage hole could be caused by this haphazard deployment. This kind of coverage gap in the network is notoriously difficult to detect. Coverage gaps could emerge as a result of the ever-changing nature of the network. Additional study is needed to forecast these kinds of network issues and identify potential solutions. The optimal computing complexity for three-dimensional coverage has thus far remained undiscovered, while the majority of present-day researchers have focused on two-dimensional space.

Anomaly detection: One of the most exciting areas of WSN research is anomaly detection, and several researchers have come up with new methods to help with this problem. When something goes wrong with a WSN, it might cause transmission delays, increased communication overhead, or even data misinterpretation by the sensor nodes. While the authors of the aforementioned works did a good job of focusing on anomaly detection, there is still a need for more research on what to do once an anomaly has been detected and how to mitigate its effects. Because anomaly detection methods vary depending on the application, picking an algorithm to use in heterogeneous WSNs is no easy feat. The detection algorithm must meet the requirements of detection speed and accuracy. In terms of routing, the majority of the currently available routing techniques were designed to gather data from a singular source and send them to a singular destination. Packet collision is a potential issue in WSNs that have numerous sources and destinations. A growing area of interest in WSN research is the development of cooperative routing protocols that do not cause collisions when there are several sources and targets. The nodes' locations in mobile WSNs are subject to constant alteration. The nodes in a WSN can move around for many reasons that are outside their control. Building routing protocols that can adapt to the ever-changing network is, thus, essential.

Data aggregation: The majority of academics are concerned with approaches for data aggregation in WSNs where the sensor nodes have a uniform data rate. When it comes to WSNs, additional investigation into sensor nodes with inconsistent data rates is necessary. With nonuniform sensor nodes, data collecting becomes more complicated in mobile WSNs. The mobile sink has the potential to enhance data collection and energy efficiency. It is difficult to schedule the mobile sink in WSNs when the data are not uniform. Consideration of energy efficiency, scalability, and low cost are the primary objectives of the efficient data aggregation process using mobile sinks.

Congestion control and avoidance: Congestion control and avoidance methods need to be resilient and incredibly resistant to internal disturbances, external stimuli, and data loss in order to keep up with the ever-changing nature of WSN operations. Because of the energy and memory limitations of WSNs, it is necessary to minimize the transmission rate between nodes and implement simple congestion management techniques at each level. When it comes to autonomous and decentralized tactics, WSNs need a way to control and avoid congestion that is quick, efficient, and effective. It is recommended that congestion control implement a self-learning strategy to adapt to the ever-changing network conditions. In order for a network to self-adopt a congestion control method, it must react by adding or removing nodes when congestion is observed. To prevent network congestion, traffic estimation methods are required to detect a quick and dynamic route change. Instead of transferring data across nodes, effective mobile agent approaches are needed to collect data from sensor nodes [42,43].

Energy harvesting: Wireless sensor networks rely on small, low-power sensor nodes dispersed across the environment and powered by batteries. The sensor nodes have limited energy resources, so a tiny, cheap, and highly efficient wireless harvesting system (WHS) is

needed to keep the network running for a long period. Wireless sensor networks (WSNs) require effective protocols for wireless energy harvesting across multiple layers. There has to be an ML-based reevaluation of the dependability of routing protocols since most current ones are energy efficient. There needs to be a system in place for large-scale sensor networks that can charge and discharge themselves in response to environmental changes. To extend the life of the network, effective methods of distributing power are required. As a result, adjusting the duty cycles requires synchronization between the physical layer's power control and the MAC layer [43,44].

Quality of service: Providing what users and applications need should be the primary focus of quality of service in WSNs. Varied needs (e.g., sensor type, data rate, traffic management methods, data kinds) and applications for WSNs necessitate varied quality-of-service criteria. Consequently, it is difficult to define the standards of quality of service for various requirements or applications. It is possible to enhance the QoS by creating efficient cross-layer protocols. Creating quality-of-service criteria for a big, diverse mobile sensor network is a formidable task.

8. Conclusions

This research looks at the fundamentals of WSN routing algorithms, including their features and how they work. We also go over the pros and cons of several methods that can be employed to enhance the efficiency of routing algorithms in WSNs. In recent years, WSNs have made great strides due to the application of ML methods. This study provides a comprehensive analysis of ML techniques. Lastly, this paper highlights the difficulties of using ML to create routing algorithms in WSNs and suggests areas that could benefit from further investigation using ML in the future. Anybody interested in ML and WSNs can benefit from this conversation. The following could be part of future studies: Machine learning methods are not suitable for large-scale deployment in low-power sensors because of the computational bottleneck and energy consumption limitations of WSNs. Distributed learning approaches, on the other hand, do not have to take the whole network's data into account; therefore, they use less power, have smaller memory footprints, and demand less processing capacity than centralized learning algorithms. Distributed cooperative learning is ideal for WSNs because it eliminates the mathematical bottleneck and enables ML-based green routing while reducing energy usage.

Author Contributions: Conceptualization, A.R.G. and A.D.P.; methodology, A.R.G. and A.D.P.; software, A.R.G. and A.D.P.; validation, A.R.G. and A.D.P.; formal analysis, A.R.G. and A.D.P.; investigation, A.R.G. and A.D.P.; resources, A.R.G. and A.D.P.; data curation, A.R.G. and A.D.P.; writing—original draft preparation, A.R.G. and A.D.P.; writing—review and editing, A.R.G. and A.D.P.; supervision, A.R.G. and A.D.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data used are made available in the present work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Shafiq, M.; Ashraf, H.; Ullah, A.; Tahira, S. Systematic Literature Review on Energy Efficient Routing Schemes in WSN Asurvey. *Mob. Netw. Appl.* 2020, 25, 882–895. [CrossRef]
- Pradeepa, K.; Anne, W.R.; Duraisamy, S. Design and implementation issues of clustering in wireless sensor networks. *Int. J. Comput. Appl.* 2012, 47, 23–28. [CrossRef]
- Wategaonkar, D.N.; Deshpande, V.S. Characterization of reliability in WSN. In Proceedings of the 2012 World Congress on In-formation and Communication Technologies, Trivandrum, India, 30 October–2 November 2012.
- 4. Low, A. Evolution of wireless sensor networks for industrial control. *Technol. Innov. Manag. Rev.* 2013, 3, 5–12. [CrossRef]

- 5. Liu, X. A survey on clustering routing protocols in wireless sensor networks. Sensors 2012, 12, 11113–11153. [CrossRef]
- 6. Jain, N.; Sinha, P.; Gupta, S.K. Clustering protocols in wireless sensor networks: A survey. *Int. J. Appl. Inf. Syst. (IJAIS)* **2013**, *5*, 41–50.
- Gupta, I.; Riordan, D.; Sampalli, S. Cluster-head election using fuzzy logic for wireless sensor networks. In Proceedings of the 3rd Annual Communication Networks and Services Research Conference (CNSR'05), Halifax, NS, Canada, 16–18 May 2005.
- Kim, J.M.; Park, S.H.; Han, Y.J.; Chung, T.M. CHEF: Cluster head election mechanism using fuzzy logic in wireless sensor networks. In Proceedings of the 2008 10th International Conference on Advanced Communication Technology, Gangwon, Republic of Korea, 17–20 February 2008; Volume 1.
- 9. Ran, G.; Zhang, H.; Gong, S. Improving on LEACH protocol of wireless sensor networks using fuzzy logic. *J. Inf. Comput. Sci.* **2010**, *7*, 767–775.
- Zhang, F.; Zhang, Q.-Y.; Sun, Z.-M. ICT2TSK: An improved clustering algorithm for WSN using a type-2 Tak-agi-Sugeno-Kang Fuzzy Logic System. In Proceedings of the 2013 IEEE Symposium on Wireless Technology & Applications (ISWTA), Kuching, Malaysia, 22–25 September 2013.
- 11. Mostafa, B.; Saad, C.; Abderrahmane, H. Fuzzy logic approach to improving Stable Election Protocol for clustered heterogeneous wireless sensor networks. *J. Theor. Appl. Inf. Technol.* **2014**, *14*, 112–116.
- 12. Bagci, H.; Yazici, A. An energy aware fuzzy approach to unequal clustering in wireless sensor networks. *Appl. Soft Comput.* 2013, 13, 1741–1749. [CrossRef]
- 13. Alaybeyoglu, A. A distributed fuzzy logic-based root selection algorithm for wireless sensor networks. *Comput. Electr. Eng.* 2015, 41, 216–225. [CrossRef]
- 14. Zahedi, Z.M.; Akbari, R.; Shokouhifar, M.; Safaei, F.; Jalali, A. Swarm intelligence based fuzzy routing protocol for clustered wireless sensor networks. *Expert Syst. Appl.* **2016**, *55*, 313–328. [CrossRef]
- 15. Logambigai, R.; Kannan, A. Fuzzy logic based unequal clustering for wireless sensor networks. *Wirel. Netw.* **2016**, *22*, 945–957. [CrossRef]
- 16. Zhang, Y.; Wang, J.; Han, D.; Wu, H.; Zhou, R. Fuzzy-logic based distributed energy-efficient clustering algorithm for wireless sensor networks. *Sensors* 2017, *17*, 1554. [CrossRef]
- Wazed, S.; Bari, A.; Jaekel, A.; Bandyopadhyay, S. Genetic algorithm-based approach for extending the lifetime of two-tiered sensor networks. In Proceedings of the 2007 2nd International Symposium on Wireless Pervasive Computing, San Juan, PR, USA, 5–7 February 2007.
- 18. Bayraklı, S.; Erdogan, S.Z. Genetic algorithm-based energy efficient clusters (gabeec) in wireless sensor networks. *Procedia Comput. Sci.* **2012**, *10*, 247–254. [CrossRef]
- 19. Hamidouche, R.; Aliouat, Z.; Gueroui, A.M. Genetic algorithm for improving the lifetime and QoS of wireless sensor networks. *Wirel. Pers. Commun.* **2018**, *101*, 2313–2348. [CrossRef]
- Abo-Zahhad, M.; Ahmed, S.M.; Sabor, N.; Sasaki, S. A new energy-efficient adaptive clustering protocol based on genetic algorithm for im-proving the lifetime and the stable period of wireless sensor networks. *Int. J. Energy Inf. Commun.* 2014, 5, 47–72.
- 21. Hussain, S.; Matin, A.W.; Islam, O. Genetic algorithm for hierarchical wireless sensor networks. J. Netw. 2007, 2, 87–97. [CrossRef]
- 22. Liu, J.-L.; Ravishankar, C.V. LEACH-GA: Genetic algorithm-based energy-efficient adaptive clustering protocol for wireless sensor networks. *Int. J. Mach. Learn. Comput.* **2011**, *1*, 79. [CrossRef]
- Mann, P.S.; Singh, S.; Kumar, A. Computational Intelligence based metaheuristic for energy efficient routing in wireless sensor networks. In Proceedings of the 2016 IEEE Congress on Evolutionary Computation (CEC), Vancouver, BC, Canada, 24–29 July 2016.
- Gemeda, K.A.; Gianini, G.; Libsie, M. An evolutionary cluster-game approach for Wireless Sensor Networks in non-collaborative settings. *Pervasive Mob. Comput.* 2017, 42, 209–225. [CrossRef]
- Yang, L.; Lu, Y.; Xiong, L.; Tao, Y.; Zhong, Y. A game theoretic approach for balancing energy consumption in clustered wireless sensor networks. *Sensors* 2017, 17, 2654. [CrossRef]
- 26. Sun, W.; Yuan, X.; Wang, J.; Li, Q.; Chen, L.; Mu, D. End-to-end data delivery reliability model for estimating and optimizing the link quality of industrial WSNs. *IEEE Trans. Autom. Sci. Eng.* **2017**, *15*, 1127–1137. [CrossRef]
- Song, X.; Wang, C.; Gao, J.; Hu, X. DLRDG: Distributed linear regression based hierarchical data gathering framework in wireless sensor net-work. *Neural Comput. Appl.* 2013, 23, 1999–2013. [CrossRef]
- Bara'a, A.A.; Khalil, E.A. A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks. *Appl.* Soft Comput. 2012, 12, 1950–1957.
- 29. Banihashemian, S.S.; Adibnia, F.; Sarram, M.A. A new range-free and storage-efficient localization algorithm using neural networks in wireless sensor networks. *Wirel. Pers. Commun.* **2018**, *98*, 1547–1568. [CrossRef]
- 30. Jadhao, P.R.; Ghonge, M.M. Energy Efficient Routing Protocols for Underwater Sensor Networks—A Survey. Energy 2015, 1–4.
- 31. Kulkarni, S.R.; Lugosi, G.; Venkatesh, S.S. Venkatesh. Learning pattern classification survey. *IEEE Trans. Inf. Theory* **1998**, *6*, 2178–2206. [CrossRef]
- 32. Lu, C.-H.; Fu, L.-C. Robust location-aware activity recognition using wireless sensor network in an attentive home. *IEEE Trans. Autom. Sci. Eng.* **2009**, *6*, 598–609.
- Branch, J.W.; Giannella, C.; Szymanski, B.; Wolff, R.; Kargupta, H. In-network outlier detection in wireless sensor networks. *Knowl. Inf. Syst.* 2013, 34, 23–54. [CrossRef]

- Somaa, F.; Adjih, C.; El Korbi, I.; Saidane, L.A. A Bayesian model for mobility prediction in wireless sensor networks. In Proceedings of the 2016 International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN), Paris, France, 22–25 November 2016.
- 35. Ghonge, M.M.; Mangrulkar, R.S.; Jawandhiya, P.M.; Goje, N. (Eds.) Future Trends in 5G and 6G: Challenges, Architecture, and Applications; CRC Press: Boca Raton, FL, USA, 2021.
- Wang, Y.; Martonosi, M.; Peh, L.-S. Predicting link quality using supervised learning in wireless sensor networks. ACM SIGMOBILE Mob. Comput. Commun. Rev. 2007, 11, 71–83. [CrossRef]
- Rooshenas, A.; Rabiee, H.R.; Movaghar, A.; Naderi, M.Y. Reducing the data transmission in wireless sensor networks using the principal component analysis. In Proceedings of the 2010 Sixth International Conference on Intelligent Sensors, Sensor Networks, and Information Processing, Brisbane, QLD, Australia, 7–10 December 2010.
- Pant, P.; Rajawat, A.S.; Goyal, S.; Potgantwar, A.; Bedi, P.; Raboaca, M.S.; Constantin, N.B.; Verma, C. AI based Technologies for International Space Station and Space Data. In Proceedings of the 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 16–17 December 2022; pp. 19–25. [CrossRef]
- 39. Ugochukwu, N.A.; Goyal, S.B.; Rajawat, A.S.; Islam, S.M.N.; He, J.; Aslam, M. An Innovative Blockchain-Based Secured Logistics Management Architecture: Utilizing an RSA Asymmetric Encryption Method. *Mathematics* **2022**, *10*, 4670. [CrossRef]
- Kathole, A.B.; Katti, J.; Dhabliya, D.; Deshpande, V.; Rajawat, A.S.; Goyal, S.B.; Raboaca, M.S.; Mihaltan, T.C.; Verma, C.; Suciu, G. Energy-Aware UAV Based on Blockchain Model Using IoE Application in 6G Network-Driven Cybertwin. *Energies* 2022, 15, 8304. [CrossRef]
- Krishnamoorthy, R.; Kumar, N.; Grebennikov, A.; Ramiah, H. A high-efficiency Ultra-Broadband mixed-mode Gan HEMT power amplifier. *IEEE Trans. Circuits Syst. II Express Briefs* 2018, 65, 1929–1933. [CrossRef]
- Vitee, N.; Ramiah, H.; Mak, P.-I.; Yin, J.; Martins, R.P. A 3.15-MW +16.0-DBM IIP3 22-DB CG inductively source degenerated balun-lna mixer with integrated transformer-based gate inductor and IM2 injection technique. *IEEE Trans. Very Large Scale Integr.* (VLSI) Syst. 2020, 28, 700–713. [CrossRef]
- 43. Eswaran, U.; Ramiah, H.; Kanesan, J. Power amplifier design methodologies for Next Generation Wireless Communications. *IETE Tech. Rev.* **2014**, *31*, 241–248. [CrossRef]
- Chong, G.; Ramiah, H.; Yin, J.; Rajendran, J.; Wong, W.R.; Mak, P.-I.; Martins, R.P. CMOS cross-coupled differential-drive rectifier in subthreshold operation for ambient RF Energy Harvesting—Model and analysis. *IEEE Trans. Circuits Syst. II Express Briefs* 2019, 66, 1942–1946. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.