Measurement 178 (2021) 108974

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Routing in wireless sensor networks using machine learning techniques: Challenges and opportunities

Padmalaya Nayak *, G.K. Swetha, Surbhi Gupta, K. Madhavi

Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India

ARTICLE INFO

Keywords:
WSNs
Artificial Intelligence
Machine Learning Techniques
Routing

ABSTRACT

Energy conservation is the primary task in Wireless Sensor Networks (WSNs) as these tiny sensor nodes are the backbone of today's Internet of Things (IoT) applications. These nodes rely exclusively on battery power to maneuver in hazardous environments. So, there is a requirement to study and design efficient, robust communication protocols to handle the challenges of the WSNs to make the network operational for a long time. Although traditional technologies solve many issues in WSNs, it may not derive an accurate mathematical model for predicting system behavior. So, some challenging tasks like routing, data fusion, localization, and object tracking are handled by low complexity mathematical models to define system behavior. In this paper, an effort has been made to provide a big outlook to the current "researchers" on machine learning techniques that have been employed to handle various issues in WSNs, and special attention has been given to routing problems.

1. Introduction

A WSN is a collection of a large number of sensor nodes, usually deployed in remote areas to monitor environmental parameters like temperature, humidity, moisture, etc. The sensor nodes are equipped with various types of sensors like acoustic, pressure, motion, image, chemical, weather, pressure, temperature, optical sensors, etc. Due to this diversity of sensor nodes, the applications of WSNs are huge in a range that starts with healthcare, military, defense, agriculture to our day to day life. Despite huge applications, WSN faces many typical challenges like limited energy sources, computational speed, memory, and limited communication bandwidth, making the sensor network degrade in performance and decreasing the network lifetime [1]. Developing different algorithms for different applications is quite a challenging task. In particular, the designer of WSNs must emphasize on various issues like data aggregation, clustering, routing, localization, fault detection, task scheduling, event tracking, etc. The various challenges and issues in WSNs are illustrated in Fig. 1. The complete description is given in section III. Among all the tasks, routing is one of the important tasks as major percentage of the energy consumption takes place while routing the data packet from the source node to the destination either through a single hop or multi-hop fashion. While routing the data, the sensor network designer must focus on all the sensor node's energy consumption issues to keep the network operating for a long time. Every routing protocol has its own characteristics and specifications based on network applications and structure.

Machine Learning (ML) is a part of Artificial Intelligence introduced in the late 1950s. Over the period, it evolved and moved towards algorithms that could computationally feasible and robust enough to handle different problems like classifications, clustering, regression, and optimization in the field of medical, engineering, and computing. ML is one of the most exciting and influential technologies in today's world. ML provides computer systems with the ability to learn automatically without human involvement and take action accordingly. It creates a model by analyzing complex data automatically, quickly, and accurately. ML has the ability to learn from the generalized structure to provide a general solution to improve system performance. Because of the diversified applications, it is applied in various scientific fields of medical, engineering, and computing like manual data entry, automatic detection of spam, medical diagnosis, image recognition, data cleansing, noise reduction [144,145], etc. Recent studies prove that ML has been applied to solve many issues in WSNs. Applying ML in WSNs not only improves the system performance but also reduces the complex tasks like reprogramming, accessing the large amount of data manually, and extracting useful information from the data. So, ML techniques are extremely helpful for fetching large amounts of data and extract useful information [2–4]. For more clarity, the requirements of Machine Learning Techniques in WSNs are briefly explained in the below

* Corresponding author.
E-mail addresses: padmalaya@griet.ac.in (P. Nayak), royal_surbhi@yahoo.com (S. Gupta).

https://doi.org/10.1016/j.measurement.2021.108974
Received 31 August 2020; Received in revised form 24 December 2020; Accepted 29 December 2020
Available online 31 January 2021

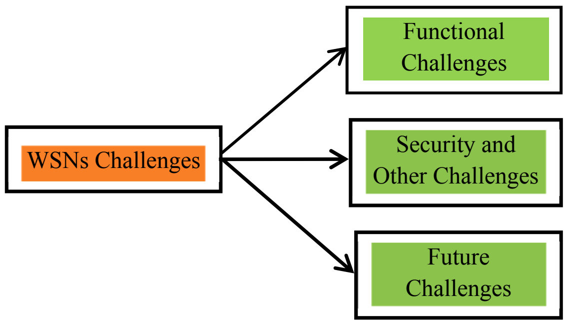

Fig. 1. Classification of WSNs Challenges.

paragraph.

1.1. Requirements of Machine Learning Techniques in WSN

(1) Energy Harvesting: It provides an advanced prediction of optimal energy consumption to be invested in performing a sensor network well in an energy-constrained environment.
(2) *Target area coverage problem*: This is another problem in WSN where ML has been focused. For the target coverage problem, ML plays a major role in finding the optimal number of sensor nodes to cover the target area.
(3) *Localization Problem*: WSNs are deployed under the water or any type of dangerous environment. The location of nodes might change due to external or internal factors. ML can help with accurate localization.
(4) *Faulty node detection*: It is assumed that sensor nodes are faulty most of the time. ML can help to detect the faulty sensor nodes and improve system performance.
(5) *Routing:* Routing plays a significant role in improving network performance by forwarding the data packet in the proper direction. Different machine learning techniques are employed to handle the dynamics of routing mechanisms
(6) *Different Levels of data abstractions*: The growing demand for WSN applications necessitate integrating WSNs with IoT, cyber-physical systems (CPS), machine to machine (M2M) communications, etc. So, intelligent decision-making systems must be developed that can be achieved through Machine Learning techniques.

1.2. Limitations of Machine Learning in WSNs

Despite several advantages of ML techniques, there are few limitations of ML techniques to apply in WSNs. The reason is WSNs always operate in a constrained environment such as limited battery power, little memory, and limited computational capacity. Learning, by example, requires large data sets of samples. WSNs consume a substantial amount of energy while predicting an accurate hypothesis and extracting the features of data samples. So, the designer of WSNs must balance the trade-off between the algorithm's computational complexity and learned model accuracy.

1.3. Search Criteria Employed

- First, good brands Journal like IEEE, Elsevier, and Springer with well-cited papers are chosen, where ML techniques are applied to WSNs.
- Second, few good "international conference papers" with high citations are referred.
- Third, we have excluded all other issues in WSNs and considered only routing issues solved by ML algorithms. It is nearly impossible to accommodate all the research papers w.r.t routing in a single review document. So, year-wise, few articles are selected, where critical routing issues are solved in WSNs. As per our knowledge, we have incorporated a considerable number of studies to develop a good survey paper.

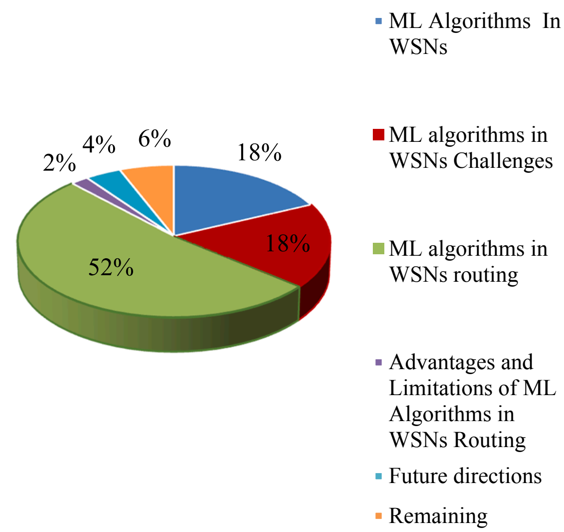

Fig. 2. Contributions of the paper.

1.4. Our Contributions

The survey in [126] focuses on Machine Learning techniques to solve various issues in WSNs, covers the period from 2002 to 2013, and the study in [137] covers the survey from 2014 to 2018. But, it is found from the existing research that most of the reviews focus on the overall issues in WSNs using ML techniques, and there is a lack of good survey explicitly presenting routing issues in WSNs by employing ML techniques.

- In our survey, we have briefly discussed ML techniques used to solve the general problems in WSNs.
- Further, we have highlighted the various issues in three categories that give a warm touch up for a better understanding of routing issues in WSNs.
- Finally, we have focused exclusively on routing issues in WSNs that have been solved by ML techniques covering the period up to 2020. In addition to this, we have discussed the advantages and limitations of ML-based routing in WSNs and discuss the open issues for future research. Our contribution has been illustrated in Fig. 2

The rest of the paper is partitioned as follows. Section 2 gives an overview of machine learning techniques used in WSNs, and Section 3 discusses the specific challenges of WSNs handled by Machine Learning techniques. Section 4 presents ML techniques used to solve routing problems in WSNs. Conclusion and Future Work is discussed in Section 5.

2. Machine learning techniques in WSNs: An brief overview

In this section, we have given an overview of various machine learning techniques based on their learning behaviors for a better understanding of ML techniques in WSNs. Broadly, it can be classified into five categories based on their learning principles. These are listed as SUPERVISED, SEMI-SUPERVISED, UNSUPERVISED, REINFORCEMENT LEARNING TECHNIQUES, and EVOLUTIONARY COMPUTING ALGORITHMS [5]. The overall classification of machine learning techniques in WSN is presented in Fig. 3. Further, the routing issues handled by

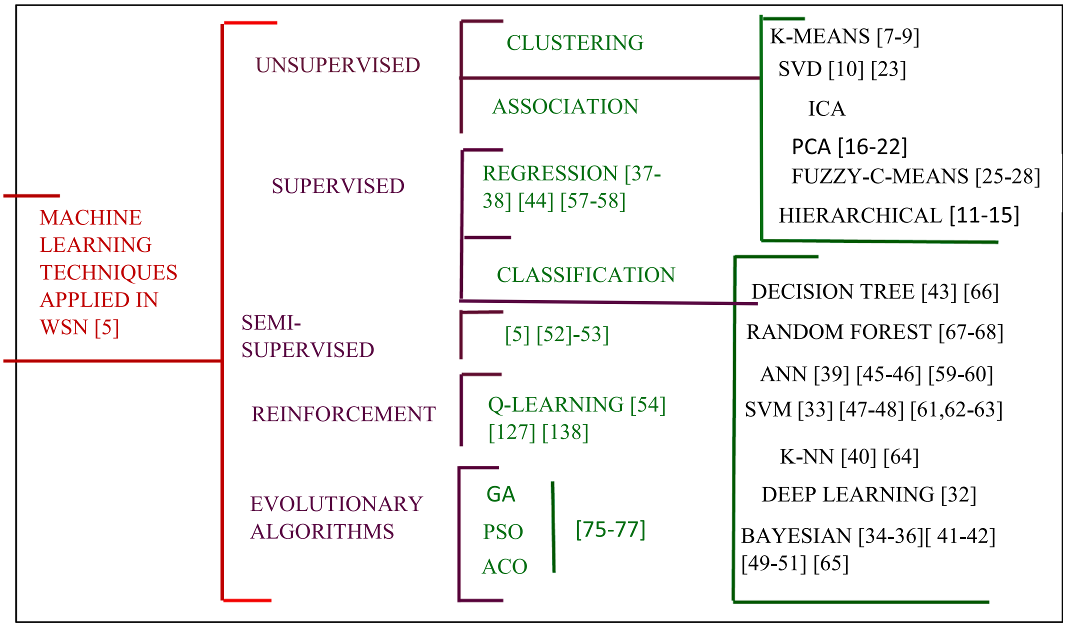

Fig. 3. Classification of Machine Learning techniques applied in WSN.

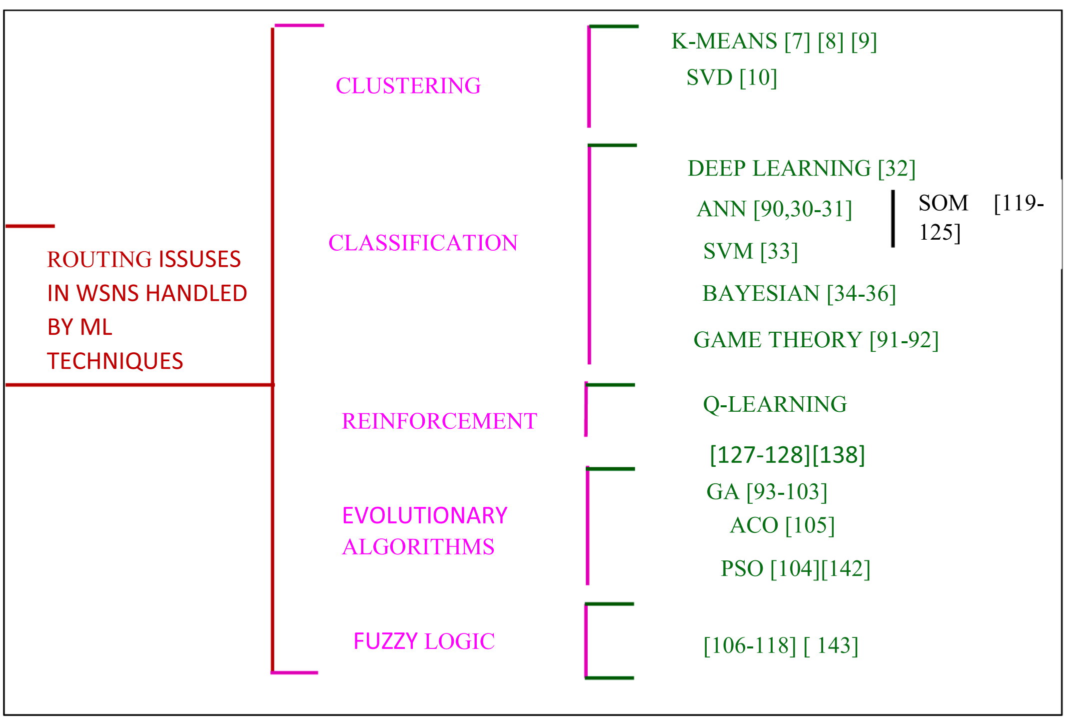

Fig. 4. Routing issues in WSNs handled by ML Techniques.

machine learning techniques are extracted from Fig. 3 and illustrated in Fig. 4. The complete description is given in Section IV for easy demonstration.

2.1. Unsupervised learning

Unsupervised learning is a training algorithm subset of AI that uses the information which is neither classified nor labeled and process without any human guidance. It can solve more complex problems than supervised learning. Unsupervised learning can figure out similar data and partition it into clusters with add on features to filter undesired data samples. Mostly unsupervised learning is used in WSNs to solve connectivity problems[6], data aggregation, clustering, routing [7–10], and other issues like anomaly detection, etc. For example, K-means clustering, hierarchical clustering, and fuzzy-c means are used for WSN clustering, and dimensionality reduction also comes under unsupervised learning that includes Principal component analysis (PCA), Independent Component Analysis (ICA), and Singular value decomposition (SVD).

2.1.1. K-means clustering

K-means clustering is an extremely popular and simplest algorithm that divides the data points into 'k' clusters or groups. A larger value of 'k' tends to smaller groups, whereas the smaller value of 'k' implies larger groups. In K-means, grouping in each cluster is identified by creating a centroid for that cluster. These centroids act as the heart of the clusters, which capture the points closest to them and include them in the cluster. The points are allocated to these clusters based on Euclidean distance to their centroids. The mean of every cluster is recomputed as new centroids, and the operation continues until the optimal cluster centroids are found. The K-Means follows the principle of minimized centroid as given in Eq. (2). K-means are mostly used in routing for selecting optimal cluster heads (CHs) in WSNs [7–9].

$$E_{k-means} = \frac{1}{C}\sum_{k=1}^{c}\sum_{x\in Q_k}\|x - C_k\|^2 \tag{1}$$

where C implies the number of clusters, Q_k identifies the K^{th} cluster, C_k defines the centroid of cluster Q_k.

2.1.2. Hierarchical clustering

In general, hierarchical clustering is meant for more massive data sets. Similar data objects form the cluster, and these clusters are arranged in hierarchical order either in a top-down or bottom-up approach. Hierarchical clustering is widely used in solving routing in WSNs, energy harvesting issues in WSNs. Hierarchical clustering is used to solve routing, data aggregation [11], routing, synchronization [12], mobile sink [13,14], and energy harvesting [15].

2.1.3. Principal Component Analysis (PCA)

Large datasets are common in many applications and difficult to interpret. Principle component analysis (PCA) belongs to an unsupervised learning algorithm that reduces the dimensionality of datasets, increases interpretability, and at the same time, preserves the data loss. PCA can compress the features, whereas K-Means can compress the data set. PCA can also be used to filter the noisy data. In WSN, PCA is applied at the individual node level, at the cluster head (CH) level to reduce the communication overheads. It also reduces the buffer overflow. Many algorithms have adopted PCA for different applications like localization [16], target tracking [20], data aggregation [17–19], and fault detection [21,22].

2.1.4. Singular Value Decomposition (SVD)

Singular value decomposition (SVD) simply displays the interesting geometrical property of PCA. It has some beautiful algebraic features which can be applied theoretically and geometrically for linear transformations in data science. It is also used in addressing routing [10] and data aggregation issues in WSNs [23].

2.1.5. Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a modified version of PCA. ICA analyses the data from various sources like business intelligence, social networking, digital images and removes the higher-order dependencies, which are not possible in PCA.

2.1.6. Fuzzy-c-means

Fuzzy-c-means clustering was developed in 1981 by Bezdek by employing a Fuzzy set theory [126]. This technique uses similarity measures such as intensity, connectivity, and distance to identify clusters. This clustering technique produces a better result in finding optimal cluster centers compared to k-means clustering when the data sets are overlapped with each other. Usually, it is used in image segmentation, business intelligence, pattern recognition, bioinformatics, etc. Mostly it is used in WSNs to solve localization [25,26], mobile sink [27], and connectivity problems [28].

2.2. Supervised learning

The term Supervised learning originates from the fact that the whole process is monitored by a supervisor. It is an influential tool to classify the data and process the data through machine learning languages. In a supervised learning algorithm, the training model comprises of known input datasets and known responses (output data). When new inputs are given, it maps an input to output based on the known input–output pairs. It is a very beneficial tool to solve classification and regression problems. For instance, Decision Trees, Neural Networks, Random Forest, Support Vector Machine (SVM), and k-nearest neighbor (K-NN) belong to supervised learning [5]. The supervised learning algorithms have been efficiently applied to solve routing problems [29–36], localization problems [44–53], event detection problems [56], target tracking [55], and sensor fusion issues in WSNs [37–43].

2.2.1. Regression

Regression techniques are primarily used for two purposes. One is for prediction, and the other is for forecasting. In some cases, the regression techniques are used to define the causal relationship between the dependent and independent variables, as expressed in Eq. (1).

$$u = f(v) + r \tag{2}$$

where *u* notifies the dependent variable, *v* implies the independent variable, *f*(*v*) builds up the relationship between *u* and *v*. *r* defines the error rate. To apply the regression technique for prediction and casual relationships, the designer must be careful to defend why the existing relationship has predictive power for a new-fangled context and how casual relationships can be defined. Many variations of regression techniques are used in the literature, such as regression with more predictive variables than observation, prediction variables measured with errors, and casual interference with regression, etc. Regression technique is used in WSNs to handle various issues like data aggregation [37,38], localization [44], connectivity issues [57,58], etc.

2.2.2. Classification

The classification algorithms learn from input data and use this learning to classify new data points. The classification algorithms may be a single class or multi-class algorithms. The algorithms like Artificial neural networks (ANN), support vector machine (SVM), and k nearest neighbor (K-NN), Bayesian learning, Random forest, decision trees efficiently solve different challenges and issues in WSNs. These are discussed in the following section briefly.

2.2.2.1. Artificial Neural Networks (ANN). An artificial neural network (ANN) is a mathematical model that imitates the human brain for performing the tasks. ANN is a huge collection of neurons that process the input data and produce the correct output. ANN consists of three layers called the input layer, one or more hidden layers, and the output layer. The input data is given to the input later and processed by the hidden layer using mathematical models, and the output layer produces accurate outputs. The ANN has been applied in WSNs to solve various issues like routing [30,31,90,119–125], node localization [45,46], data aggregation [39], congestion control [59,60], etc.

2.2.2.2. Support Vector Machine (SVM). Support Vector Machine is a category of supervised ML technique that gives the best classification from a given data set by using a hyperplane by coordinating individual observations. SVM can solve both linear and no-linear problems and most suitable for large datasets. SVM is applied to WSNs to solve various issues like routing [33], localization [47,48], fault detection [22,61], congestion control [62], and connectivity issues [63].

2.2.2.3. K-Nearest Neighbor (k-NN). K-Nearest neighbor (k-NN) is one of the most popular, straight forward instance-based learning methods used to solve regression and classification problems. k-NN mostly considers the distance between the given training sample and the test sample. The various distances like Hamming distance, Euclidean distance, Manhattan distance and, Chebychev distance function are considered in k-NN. This method finds the missing samples from the featured space and decreases the dimensions. The k-NN has been applied for data aggregation [40] and anomaly detection [64] in WSN applications.

2.2.2.4. Deep learning. Deep learning is a data learning method with a multi-layer perception that belongs to the ANN family. It mimics the communication and information processing systems of the human brain and processes the data for detecting objects, translating languages, recognizing speeches, and making decisions. Deep learning is used to handle many issues in WSNs like anomaly and fault detection, energy harvesting, data quality estimation, and routing [32].

2.2.2.5. Bayesian. Bayesian learning is a statistical learning approach, finds the relationship between the datasets by learning conditional independence from different statistical methods. Bayesian learning takes different prior probability functions and new information to determine posterior probabilities. If a set of inputs are represented by Y_1, Y_2, $Y_3 \ldots Y_n$, and returns, a label θ, the probability of p(θ) must be maximized. Several issues in WSNs such as routing [34–36], data localization [49–51], aggregation [41,42], fault detection, connectivity, and coverage problems [65] have been solved by Bayesian learning methods.

2.2.2.6. Decision trees. Decision trees (DT) belongs to the supervised learning ML techniques that use sets of if then else rules to enhance the readability. DT contains two types of trees. One is the leaf node, and another is the decision nodes. DT predicts a class or target based on the decision rules and creates a training model inferred from training data. There are many advantages of decision trees like transparency, less ambiguity in decision making, and allows for a comprehensive analysis. Decision trees are applied in WSNs to handle various issues like connectivity [66], data aggregation [43], mobile sink, etc.

2.2.2.7. Random forest. Random Forest (RF) Algorithm is a supervised learning algorithm having a collection of trees, and each tree gives a classification. RF works on the two principles; first, it creates a forest classifier, then produces the results. RF works well for heterogeneous data with a vast number of data sets. RF is used in WSNs to solve problems like MAC protocols [67] and sensor network coverage [68].

2.3. Semi-supervised learning

Any machine learning algorithm requires training data to learn from it. Semi-supervised learning uses both known and unknown data sets for training and predicts the output based on the trained data. In semi-supervised learning, the data is first clustered using unsupervised learning; later, the remaining data is labeled using supervised learning [5]. It's relatively expensive to gather input and output pair training data in a semi-supervised learning algorithm in practical applications. Semi-supervised learning is applied in WSNs to solve various issues like data aggregation, localization [52,53], and fault detection in WSNs.

2.4. Reinforcement learning

Reinforcement Learning is one category of machine learning that learns from the environment in the absence of a training dataset. It tries to take suitable action to maximize the reward points according to the situation. The Q-learning and deep Q-learning is the example of reinforcement learning [5]. Reinforcement learning is used to solve routing issues efficiently in WSNs [54,127,138].

Table 1
Functional Challenges in WSN.

SL. No	WSN Challenges	Machine learning technique	Studies
1	Clustering and data aggregation	PCA	[17–19]
		Regression	[37,38]
2	Event detection and Query Processing	KNN	[56,69]
		Deep learning	[70]
		PCA	[71]
3	Energy harvesting	Deep learning	[72]
		Reinforcement learning	[73,74]
4	Coverage and Connectivity	Fuzzy c means	[28]
		Regression	[57,58]
5	Localization and Object Tracking	ANN	[45,46]
		SVM	[47,48]
		Semi supervised	[52,53]
6	Mobile Sink	Evolutionary algorithm	[2,27,75–77,139,140,146]
7	Congestion Control	SVM	[62]
		ANN	[59,60]
		Reinforcement Learning	[78,79]
8	Routing	ANN	[119–125]
		Evolutionary algorithms	[93–105]
		Fuzzy logic	[106–118]

2.5. Evolutionary computing algorithms

Evolutionary algorithms are a subcategory of artificial intelligence (AI) that uses a heuristic-based approach to solve problems that can not be solved by polynomial time. Evolution algorithms are motivated by nature and mostly used to solve optimization problems. These algorithms include genetic algorithms, particle swarm optimization, ant colony optimization, etc. [5]. Evolutionary algorithms are used to handle various issues and challenges in WSNs efficiently [75–77].

3. Challenges in WSNs and machine learning techniques

Usually, sensor nodes are deployed in hazardous environments where we leave the network run automatically without any human intervention. The designer of WSNs must consider the limited battery power, memory constraints, link failure, dynamic changes in topology (sometimes), and decentralized control. In this section, we discuss various challenges in Wireless Sensor Networks that are handled by machine learning techniques. We have classified WSN challenges into 3 types. These are; (i) Functional challenges, (ii) Security and other challenges, (iii) Future challenges [126]. These challenges are discussed in each subsection, and we have extracted the features and represented them in a tabular form for the clarity of the demonstration.

3.1. Functional challenges

There are many challenges like Clustering and Data aggregation, event detection and query processing, energy harvesting, MAC management, Localization, Object tracking, Mobile sink, Congestion control, Coverage, and Connectivity, Routing issues handled by different machine learning algorithms in WSNs. All these challenges are categorized as functional challenges. Table 1 summarizes the functional challenges of WSNs.

3.1.1. Clustering and data aggregation
In large scale networks, it is a mandatory requirement in WSNs that the sensed data must be delivered to the sink node directly. Clustering can help to transmit the data directly to the sink node and saves a tremendous amount of energy. Efficient cluster head selection is another challenging task in clustering that leads to minimal energy consumption. CH collects the data from other sensor nodes within the cluster, performs

data aggregation, detects the faulty nodes, and removes the faulty nodes from the network [7–9,17–19,37,38].

3.1.2. Event detection and query processing

Many applications of WSNs require that event should be detected around the moving objects and should be delivered to the user. Furthermore, the events occur at different locations and last for a longer period that is unknown in advance. WSNs can be monitored in three ways, such as event-driven, continuous, or query-driven. ML helps to provide effective query processing and solutions to query processing, to detect events, assess event validity with limited resource facility. Although event detection and query processing have taken huge attention from the research community [56,69–71], there is still a lack. It demands developing advanced event detection and query processing technique applying different machine learning techniques.

3.1.3. Energy harvesting

Energy harvesting has appeared as an alternative for providing sensor nodes power to operate for a longer time in an open environment. It is the process of transferring ambient energy to electrical energy. Several challenges might pose while successfully transporting energy harvesting technology into WSNs [15,72–74].

3.1.4. Localization and object tracking

Localization is the process of locating the sensor node's geographic position because most of the WSN application is based on location [45-48,52,53]. In large scale networks, it is hard to fix up the Global Positioning System (GPS) hardware in each sensor node as it is not financially feasible. Further, GPS does not support the indoor environment.

3.1.5. Mobile sink

In WSNs, sensor nodes collect the information and send the information to the sink node either through a single-hop or multi-hop manner. The node near to BS becomes the bottleneck and creates a hotspot problem. So, the concept of Mobile sinks has been introduced in WSNs research. The mobile sink gathers the information by moving from one sensor node to the other. As it is difficult for a mobile sink to visit each sensor node, scheduling the mobile sink is a focused research issue. Instead of visiting all the sensor nodes, the mobile sink collects the information from Rendezvous points (RPs). Other sensor nodes send the information to RPs. Sometimes multiple mobile sinks are used to avoid delay, but it is too cost-effective. ML can help to find optimal RPs and avoid delays by using a mobile sink [2,27,75–77,139,140,146].

3.1.6. Congestion control

Usually, congestion occurs only when the volume of information crosses the capacity of the communication channel. In the context of WSNs, congestion occurs when the communication channel transmits more data compared to the size of the bandwidth. Congestion affects the end to end delay, packet loss, QoS, and overall energy consumption. Mostly, congestion occurs at the node level or the link level. Congestion occurs at the node level due to the packet arrival rates, and Link level congestion occurs due to a lower bit error transmission rate between two nodes and collision. ML algorithms can estimate the traffic accurately, reduce the end to end delay, adjust the transmission range dynamically [59,60,62,78,79].

3.1.7. Coverage and connectivity

Coverage and connectivity is a major issue in WSNs. Mostly, WSNs are deployed randomly or deterministically in a particular area. Coverage means how efficiently a sensor node monitors the specified area. Connectivity means a sensor node should be able to reach the BS station directly or through relay nodes. If the sensor nodes do not cover the area, there will be a gap in between nodes. ML can help to estimate the optimal number of sensors to cover the target area and dynamically change the routing path if any connectivity issue occurs [28,57,58].

Table 2
Security and other Challenges in WSN.

S. No	WSN Challenges	Machine learning technique	References
1	Security and anomaly detection	Regression	[80,81]
		SVM	[82,83]
		Decision tree	[84]
2	QoS management	Reinforcement Learning	[78,79]
		ANN	[59,60]
3	Link Quality Management	Matric-map	[85,86]
4	Resource allocation and task scheduling	Fuzzy logic	[87]
		Evolutionary Algorithm	[88]
		ANN	[89]
5	Fault detection	PCA	[21,22]
		SVM	[22,61]
		Bayesian	[65]

3.1.8. Routing

Routing is one of the primary issues in WSNs due to the diversified applications. ML helps to find an optimal route by consuming less energy while transmitting the packet from the source node to the sink node [93–125]. By doing so, it extends the network lifetime. The details are explained in a separate section (Section 4).

3.2. Security and other challenges

This section discusses the security and other challenges in WSNs in detail. These challenges include anomaly detection, QoS management techniques, Link quality management, resource allocation and task scheduling, fault detection, etc. For clarity of the demonstration, we have summarized in a tabular form, as shown in Table 2.

3.2.1. Anomaly detection

Anomaly detection is one of the significant concerns in WSNs. Anomalies in WSNs lead to an end to end delay, inaccurate sensor readings, transmission overheads, etc. So. various techniques have been developed to detect anomalies and to protect from multiple attacks such as black hole attack, gray hole attacks, wormhole attacks, and hybrid anomalies [80–84].

3.2.2. QoS management

Quality of Service (QoS) ensures high priority in delivering real-time data to the destination. In the context of WSNs, it suffers from bandwidth and energy constraints in the timely delivery of the data at the destination. Most of the time, it is assumed that sensor nodes are faulty. So, data aggregation, query processing, unbalanced traffic, data redundancy, scalability, along with randomly deployed sensor nodes, poses enormous challenges to QoS requirements in WSNs [78,79]. The network-specific, as well as application-specific QoS requirements in WSNs, are well managed by ML techniques listed in [59,60]

3.2.3. Link quality management

Link quality estimation is an essential feature in WSNs as it depends upon the environmental parameters like signal quality, interference, etc. In [85,86], a metric map is developed to measure the link quality using supervised learning. This study uses received signal strength indicator (RSSI), communication channel load, buffer size, and forward–backward probabilities. More such ML techniques must be investigated for accurate link quality measurement.

3.2.4. Resource allocation and task scheduling

Energy-saving is the major issue in WSNs. This goal can be achieved either by developing suitable communication protocols or monitoring the activities of sensor nodes. For instance, sensor nodes waste their energy by listening to other node's transmission [87–89]. Such type of active operations can consume more energy. ML techniques can be configured in the sensor nodes to optimize resource allocation and

Table 3
Future Challenges in WSN.

SL. No	WSN Challenges	Machine learning technique	References
1	Compressive Sensing and Sparse Coding	Bayesian, ICA, Dictionary Learning, SVD	[135]
2	Detection of data Spatial and Temporal Correlations	Need to be explored	[129,130]
3	Distributed and Adaptive Machine Learning Techniques	Need to be explored	[131–134]
4.	Resource Management	Need to be explored	[136]

Table 4
Major issues and challenges in WSN handled by ML Techniques.

SL. No	WSN Challenges	Machine learning technique	Remarks
1	Clustering and data aggregation	PCA [17]	Improved Data aggregation process
		Regression [38]	Improved network lifetime
2	Localization and Object Tracking	SVM [48]	Improved localization accuracy
		Semi-supervised [52]	Minimized error rate
3	Routing	ANN [119]	Reduced energy consumption of sensor nodes and prolong the network lifetime
		Evolutionary algorithms [102]	Path optimization
		Fuzzy logic [117]	Increased network lifetime
4	Mobile Sink	Evolutionary algorithms [2,76,139]	Optimal mobile sink path selection
		[140]	Optimizes communication distance, reduces energy consumption by avoiding long-distance communication, and reduced energy consumption
5	Security and anomaly detection	Regression [81]	High detection rate
		SVM [82]	Minimizes the complexity of training and testing phases
6	QoS management	Reinforcement learning [79]	Achieved QoS by maintaining Reliable topology
		ANN [60]	Achieved QoS by avoiding congestion

power management, etc.

3.2.5. Fault detection

Usually, WSNs are deployed in hostile environments where the human attendant is not feasible. The fault in WSNs occurs due to several reasons such as battery failure, communication link failure, node failure, software failure, topology change, dynamic environment, etc. So, the detection of faulty nodes as well as other faults in WSNs is a quite challenging task. Several fault detection mechanisms are discussed in the literature [21,22]. The use of machine learning techniques in WSNs reduces the complexity and increases the accuracy [22,61,65].

3.3. Future challenges

There are many challenges like comprehensive sensing, detecting data spatial and temporal correlation, proper resource management, distributed and adaptive machine learning techniques must be developed for in-network processing of data instead of exhausting the nodes for high computational tasks. We have referred to a few studies that discuss these issues [129–136] and summarized in Table 3. ML-based Data compression and dimensional techniques reduction can be used to compress the data instead of traditional compressed techniques to produce a better energy-saving scheme. Data correlation is an important issue that needs to be addressed in hierarchical clustering. Only one node must be active at a particular time in a cluster to monitor the cluster area to save energy. In WSNs, most of the time, sensor nodes exhaust energy by over listening to other nodes. The sensor nodes equipped with machine learning techniques can be able to manage the resources and power allocation schemes efficiently.

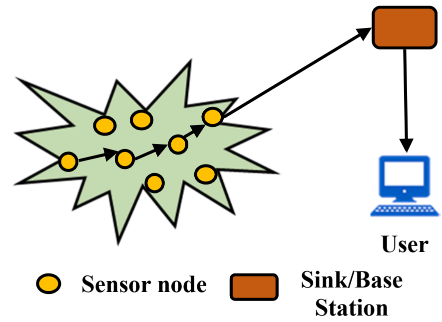

Fig. 5. Example: Routing in WSNs.

To avoid large-size tables, out of many challenges and issues, we have selected a few significant problems and challenges that have been handled by ML techniques in WSNs and summarized in Table 4. Already previous sections present some hints about these issues, and the remark column of Table 4 describes the functionalities of each challenge. For example, Clustering and data aggregation [17] issues handled by ML techniques can improve the efficiency of the data aggregation process and prolongs the network lifetime. Similarly, ML-based localization and object tracking methods minimizes the error rate and improves the localization accuracy. Routing [24,102,117], Mobile sink [2,76,139,140], Security and anomaly detection [81,82], and QoS management [60,79] are the major issues that have been handled by ML techniques, and we limit the discussion here to avoid duplicity.

4. Routing in WSNs using machine learning techniques

This section discusses various machine learning techniques, how efficiently ML has been applied to handle WSNs routing. As the applications of machine learning are widely applied in different aspects of WSNs, in this section, we have focused on only the routing issues that have been handled by ML techniques. In WSNs, collecting sensed information, again extracting useful information from the gathered data, processing the data, delivering the data to the BS in an energy-efficient manner, and enhancing the network lifetime are the key issues. So, energy conservation is one of the critical design goals in WSNs, and routing protocols are the best-known solutions for energy conservation. Large scale networks undoubtedly present a large amount of data to be transmitted, processed, and received. It is nearly impossible to transfer all the data to the BS due to sensor limited constraints and bandwidth constraints. Opting machine learning techniques in routing can process a massive amount of data with less amount of time and provides accuracy. To handle these issues, many routing protocols are developed in the recent past. For the sake of simplicity, we have depicted the routing issues dealt with by ML techniques in Fig. 4 in section II, and we have demonstrated a basic routing example in Fig. 5.

4.1. Advantages of ML techniques in WSNs routing

The main advantages of ML-based routing are listed as follows.

- ML does not require reprogramming due to the environmental changes.
- ML also reduces communication overhead as well as delay.
- ML also helps to select the optimal number of cluster heads in routing.
- ML reduces the complexity of routing and satisfies the QoS requirements using simple computational methods and classifiers.

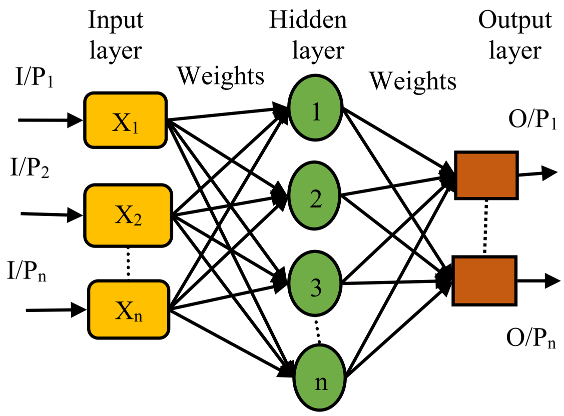

Fig. 6. A simple artificial neural network.

4.2. Limitations of ML techniques in WSNs routing

Despite many advantages, machine learning algorithms have few limitations for applying in WSNs routing.

- A large amount of data consumes a large amount of energy while processing the data. So, there is always a trade-off between energy constraints and high computational complexities involved in ML techniques.
- Again, the system's performance depends upon the past data, which is quite tough to acquire in WSNs.
- In a real-time scenario, validating the results through a machine learning algorithm is quite a difficult task. Sometimes, identifying machine learning algorithms to solve a particular routing issue is quite tricky.

4.3. Data routing using ANN

Artificial Neural Network (ANN) is an information processing system or computational model inspired by the biological structure of neurons. ANN is the basic building block of Artificial Intelligence and problem-solving that could be nearly impossible for a human. The human brain consists of billions of cells, and each cell is called a neuron. The interconnection of the number of neurons having self-learning capabilities produces better results by processing a large amount of data from the outside world. In general, ANN consists of three layers. These layers are:

- Input layer
- The intermediate layer (Hidden Layer)
- Output layer

The hidden layer can be a single layer or multiple layers that depend on the applications. The basic structure of the simplified ANN is depicted in Fig. 6. ANN involves a large number of processors operating in parallel and arranged hierarchically. The input layer processes raw data and passes it to the other layer. The hidden layers take the data from the preceding layer rather than the raw data and process it and send it to the next layer, similar to the optic nerves in human visual processing. Finally, the output layer produces the output. In [90] ELDC, a robust routing protocol is discussed based on ANN that trains the protocol using various parameters distance between the sensor nodes, CHs, residual energy, broader nodes, and base station. ELDC uses a backpropagation neural network (BPNN) to elect some reliable CHs that balance energy consumption and avoids data loss in WSNs. ANN operates on two types of learning procedures. One follows the principle of self-learning/competitive learning, and another follows the corrective learning principle. Self-Organizing Map (SOM) is the most widespread ANN model and belongs to a competitive learning model that means human intervention is not required during the learning process. It should have little knowledge about the physical characteristics of the input data.

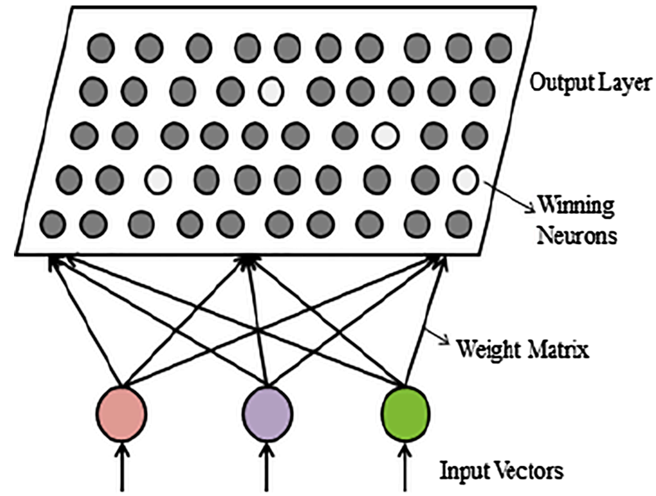

Fig. 7. Example of SOM.

4.3.1. Self-Organizing Map (SOM)

A Self-organising map (SOM) is one category of (ANN) which is trained by unsupervised learning to reduce the dimensionality (see Fig. 7). The SOM is developed by professor Kohonen in 1980 and recognized as Kohonen maps. The SOM is used mostly for clustering by grouping similar data points. SOM is widely used for solving routing problems in WSNs.

In [119], the author has developed a routing protocol called Energy Based clustering Self Organisation map (EBCS) that uses energy level and coordinates of nodes as the parameters to find CHs. The EBCS forms energy balanced clusters that minimize the energy consumption of nodes by balancing the load in a network, which leads to network lifetime extension. In [120], the author uses a neural network to build efficient topologies in WSNs. These Efficient topologies make the routing procedure easier and reduce the energy consumption efficiently in WSNs. Authors of [121] use ANN for routing purposes. The main advantage of using ANN is to speed up the lookup table process. The size of the lookup table is not influenced by the decision of the speed such as where to send the packet next. The research in [122] presents an Adaptive Resonance Theory (ART) neural network for routing in WSNs. This work improves the lifetime of the network by using some mechanisms that include minimum CH separation distance, a CH rotation system, ART1 based CH election, and load balancing cost functions. The research work in [123] has proposed an improved PEGASIS routing algorithm based on neural network and ant colony algorithm. This protocol focuses on the neural network for CH selection based on the location, residual energy of node, and neighbor node to the base station. In [124], a solar energy prediction model is proposed based on neural networks and proved its energy efficiency. It is concluded that nodes having higher residual energy are having more robust energy harvesting capacity.

An enhanced NN based RZ LEACH protocol is proposed in [125] that uses hybrid ACO/PSO based routing to enhance the network lifetime in WSNs. NN is used to emphasize the cluster head selection process.

4.4. Bayesian-based routing

Few Naïve Bayes routing protocols for WSNs are discussed in [34–36]. In [34], the research work focuses on the selection of CHs for routing and prolongs the network lifetime by reducing energy consumption. Even Naïve Bayes has an additional property that ensures of adding and removing new features dynamically. In [35], a data

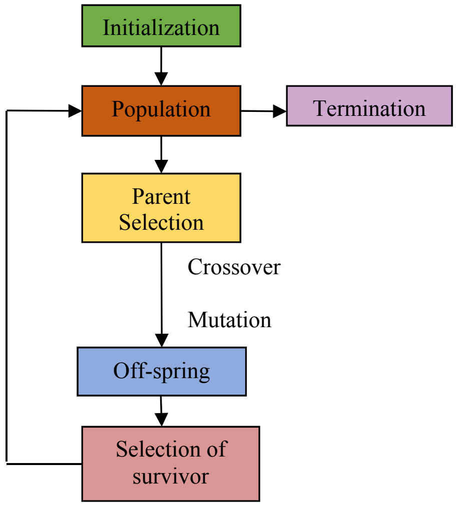

Fig. 8. Basic Principle of Genetic Algorithm.

collection technique is discussed based on the Bayesian approach that emphasizes creating a projection vector in each reiteration of routing by incorporating a target node selection. A Bayesian learning method is used in [36] that is suitable for centralized as well as a decentralized infrastructure. This method also uses a scheduling approach to balance energy consumption.

4.5. Data routing using K-means

In [7], a well-known clustered based routing is discussed that finds the optimal number of clusters for WSNs routing. This routing protocol improves the packet delivery ratio (PDR), throughput, minimize energy consumption, and reduces communication overhead. In [8], the author discusses a protocol called as EECPK-means, takes a random initial center point, and finds the optimal center point in the cluster. The optimal number of CHs is found based on Euclidean distance and residual energy of the sensor nodes. EECPK-means follows a multi-hop path from CHs to BS by maintaining energy consumption uniformly and protects data loss. An efficient K-means (EKMT) is discussed in [9] that finds optimal cluster heads close to the member nodes of the cluster and sink node. This technique selects the CHs dynamically and enhances the network lifetime, reduces the delay, and produces better throughput.

4.6. Routing using reinforcement learning

In [138], the author has explored the applications of the Q-learning algorithm to solve multicast routing problems in wireless ad hoc networks through reliable resource allocation. In [127], the author has used unlicensed ultra-wideband (UWB) technology based on the principle of Q-learning [138] to enhance the performance in geographic-based routing. The sensor node energy and delay are used as the metrics to formulate the learning reward function in Reinforcement learning-based geographic routing protocol. In this routing protocol, UWB devices are fitted with the CHs to detect the locations of the nodes. Each node is equipped with a routing table that maintains the information about neighbour nodes to exchange the data by sending a simple “hello” messages and in this way the routing path is established. The main advantages of reinforcement learning are (i) it does not require global information to achieve acceptable routing solutions. (ii) the UWB communication operates on a constant frequency band (3.1–10.6 GHz) provided by the Federal Communications Commission (FCC) [128]. The author has discussed an enhanced geographic routing for WSNs called “Q-Probabilistic Routing” (Q-PR) that finds optimal paths by learning from the previous routing based on reinforcement learning and Bayesian decision model.

4.7. Routing using computational evolution algorithm

4.7.1. Genetic Algorithm (GA)

Genetic Algorithm is a bio-inspired algorithm evolved from Charles Darwin’s theory of natural selection. Mainly, GA operates on the principle of natural selection, where the fittest individuals are selected to reproduce offspring for the next generation. To generate new offspring, GA uses the following biological operators. The principle of GA is depicted in Fig. 8.

- Selection operator
- Crossover operator
- Mutation operator

Each generation consists of a set of populations, and each population is characterized by a set of parameters called genes. The size of the population can be selected based on a specific problem. The fitness function determines the fitness value of an individual that is the skill of an individual to compete with other individuals. The selection operation selects the fittest individual for the next generation. The two best fittest chromosomes are called parents. The parents are mated at a crossover point to produce new offspring. Some of the genes of the new offspring are muted with a low random probability, and the process continues till no new generation can be created. Various researchers have incorporated GA in WSNs to solve routing problems. In [93], GA is used to cluster the non-uniformly distributed nodes and select the cluster heads (CHs) among the clusters. The algorithm can also be used to restructure the clusters and cluster heads when geographic location changes or failure of CH occurs. In [94], the author has discussed a routing protocol for energy harvesting in WSNs. The main purpose of energy harvesting in WSNs is to maximize the operating environment of WSNs instead of maximizing the network lifetime why because the sensor nodes rely on the energy harvested from the environment rather than batteries. In [95], the author has introduced sensor mobility to minimize the distance between CHs and base stations. Obtaining this procedure, the author has proved the energy efficiency over the LEACH routing protocol. A Routing protocol for a two-tier sensor network based on GA is proposed in [96] that expands the network lifetime by reducing energy consumption. The research in [97] discusses a flat-based routing protocol that finds an alternate path in a WSN using MOGA. The study in [98] presents a technique to improve the lifetime of a multi-sensor network using optimal traffic distribution. In [99], the work has been focused on distributed GA that proves its energy efficiency based on the required detection probability. In [100], the author has concentrated on the QoS parameters by using GA based technique. In [101], a GA based approach is used in routing for path optimization. The research in [102,103] focus on GA based techniques, where the emphasis is given on fitness function that is calculated based on two significant parameters distance and energy, and proved as energy efficient. In [146], the movement of the mobile sink is scheduled efficiently through genetic algorithm programming to maximize the Wireless Sensor Network lifetime.

4.7.2. Ant Colony Optimization (ACO)/Particle Swarm Optimization (PSO)

The research work in [104] presents an efficient routing protocol that considers two metrics, such as the transmission range of a sensor node and data forwarding load. This research focuses on the clustering method based on particle swarm optimization. In [105], the research work also follows ant colony optimization (ACO) for routing the data packet considering various parameters such as transmission distance, residual energy and finds the shortest path from the source node to the destination by consuming minimal energy.

4.8. Routing using support vector machine

In [10], the author has used a shallow light tree (SLT) along with

Table 5
Summary of SOM based Routing in WSNs.

Studies	Authors	Criteria	Network type	SOM parameters	Simulation Tool	Performance metrics
[119]	Neda Enami et al.	clustering	Homogenous WSN with static sensor nodes	Energy and coordinates of sensor nodes	MATLAB	First node dies, Half node dies, Last node dies, and Network coverage
[120]	Chiranjib Patra et al.	Design of Energy efficient topologies	WSN with random deployment of sensor nodes	Coordinates of sensor nodes	MATLAB	Node usage frequency throughout the simulation
[121]	Michal Turčaník et al	Designing routing table	WSN with a mesh topology	Node address and interface Status	MATLAB	Delay, No. of look up tables, and slices
[122]	Mohit Mittal and Krishna Kumar	Network lifetime extension	Homogenous WSNs with static sensor nodes	Sensor node positions	MATLAB	Network lifetime and no. of transmitted packets
[123]	Tao Li et al.	Optimal chain path and chain head selection	Homogenous WSN with static sensor nodes	Sensor node positions, energy, and number of neighbors	NA	NA
[124]	Junling Li, Danpu Liu	CH selection	Energy harvested WSN	Residual energy and harvested energy of sensor nodes	OMNeT++ 4.5	No. of awake nodes, Residual energy, and Throughput
[125]	Deepshikha et al.	Network lifetime extension	WSNs with mobile BS	Energy, number of neighbors, distance	MATLAB	No. of dead nodes, Remaining energy, No of packets transmitted to CH and BS

singular value decomposition (SVD) to route the data packet to the base station by considering an arbitrary topology. This network is most suitable for smart cities, specifically for health monitoring using IoT devices. But the drawback of this system is transmission overhead increases proportional to the increased number of sensor nodes. A secure cluster-based routing protocol is proposed in [33] that selects the CHs based on residual energy and distance and enhances network lifetime.

Table 6
Summary of Evolutionary computing-based routing in WSNs.

Studies	Authors	Criteria	Network type	Fitness function	Simulation Tool	Performance metrics
[93]	Zhou Ruyan et al.	Selection of Cluster Head	WSN with mobile sensor nodes	Fitness function based on cluster centers	NA	NA
[94]	Yin Wu et al.	CH selection and finding optimal routing path among each CHs	WSN where sensor nodes are energy harvested	Fitness function based on distance	OMNET++	Packet loss and energy consumption
[95]	Omar Banimelhem et al.	Mobility at the sensor nodes to reduce the distance between the sensor node and BS	WSN with mobile sensor nodes	Fitness based on distance	MATLAB	First node dies, half node dies, Last node dies, and Remaining energy
[96]	Ataul Bari et al.	Finding optimal path among each relay nodes	Two tired sensor networks	Fitness function based on energy	MATLAB	Network lifetime
[97]	Srinivas N et al.	Optimal Traffic distribution Technique	sensor nodes and BS are static	Fitness function based on distance	NA	Chi-square-like deviation form distribution
[98]	Y. Pan and X. Liu	Detection Probability	Multi-Sensor Network	Fitness function based on network lifetime and throughput constraint	NA	Network lifetime extension
[99]	Q. Qiu et al..	Emphasis on QoS parameters	Clustered WSN	Fitness function based on remaining energy	C++ based software program	Network lifetime and Energy
[100]	Navrati Saxena et al..	Emphasis on QoS Parameters	WSN with real-time data	Fitness function based on distance	NA	Delay, Energy consumption, and Throughput
[101]	Jin. M. Zhou, and A.S. Wu	Path optimization	Clustered WSN	Fitness function based on distance	NA	Clustering with Different BS positions and scalability
[102]	P. Nayak et. al	Path Optimization	Hierarchical sensor network	Fitness function based on distance and energy	NetSim Simulator	Network lifetime and No of packets sent to BS
[103]	Veena Trivedi and P. Nayak	Path Optimization	MANET	Fitness function based on energy and distance	ns-2 Simulator	Delay, Throughput, and PDR
[104]	P. Kulla, P.K. Jana	Path Optimization (PSO)	Homogenous WSN	Fitness function based on transmission range and data forwarding load	MATLAB	Network lifetime, Energy consumption, Dead sensor nodes, and No of packets sent to BS
[105]	Y. Sun, W. Dong, Y. Chen	Path Optimization (ACO)	Static WSN	Fitness function based on distance, energy, and transmission path	NA	Energy consumption and Network lifetime
[141]	M. Khabiri, A. Ghaffari	Selecting optimal cluster heads (Cuckoo Optimization)	WSN with static nodes	Fitness function based on distance and energy	MATLAB	No. of alive nodes, Minimum energy consumption, and No. of packets sent to BS
[142]	S. Tabibi, A. Ghaffari	Finding optimal rendezvous points (PSO)	WSN with mobile sink	Weight value based on the number of packets from another node	MATLAB	No. of hops, Packet Loss Ratio, Throughput, and Energy consumption
[146]	Jinghui Zhong, Zhixing Huang, Liang Feng, Wan Du, Ying Li	Scheduling mobile sink movements	WSN with mobile sink	Fitness function based on training networks	NA	Average network lifetime and Average decision time

Table 7
Summary of Various Machine Learning Techniques in WSNs Routing.

Sl. No	Machine learning approach	Studies	Topology type	Mode of operation	Node Status	QoS Status	Network Performance	Performance metrics
1.	ANN	[90]	Tree	Centralized	Static	No	Increases the data delivery ratio	Network lifetime and Energy consumption
		[30]	Tree	Distributed	Static	Yes	Balances the energy consumption and avoid data loss	Network lifetime and PDR
		[31]	Tree	Distributed	Mobile	Yes	Enhances Network Lifetime	Network lifetime, Energy consumption, and Packet loss
2.	Deep Learning	[32]	Hybrid	Centralized	Mobile	No	Overcomes congestion, packet loss, and better power management	PDR, Route discovery speed and Connectivity
3.	SVM	[33]	Hybrid	Distributed	Static	No	Extends Network Lifetime	PDR, and Energy
4.	Game Theory	[91] [92]	Hybrid	Distributed	Mobile	yes	Improves Network Lifetime Extends Network lifetime	Transmission cost The First node dies, the Last node dies, Residual Energy, and No of packets sent to BS
5.	Bayesian	[34]	Tree	Distributed	Static	No	Balances energy consumption	Number of alive nodes
		[35]	Hybrid	Centralized	Static	No	Path Optimization	Reconstruction error, Energy cost, and communication complexity
		[36]	Hybrid	Distributed, CentralizedBoth	Mobile	No	Balances the energy consumption	Energy Consumption w.r.t time
6	K-Means	[7]	Hybrid	Distributed	Static	Yes	Better PDR, Throughput, minimization energy consumption, control traffic overhead	Number of clusters, Delay, Packet Delivery Ratio, and Throughput
		[8]	Tree	Distributed	Static	No	Avoids data loss, balances the energy consumption	Network lifetime and Energy consumption
		[9]	Hybrid	Centralized	Static	No	Improves throughput and avoids delay	Throughput and delay
7.	SVD	[10]	Arbitrary	Distributed	Static	No	Suitable for IoT application	Transmission cost
8.	Q-MAP multicast/ Reinforcement Learning	[138]	Flat	Distributed	Static	No	Reduces overhead of route searching	No of packets sent, Packet delivery ratio, and Delay

4.9. Fuzzy logic-based routing protocol

The fuzzy logic (FL) was invented by Lotfi Zadeh at the University of California in 1965. Fuzzy logic computing is not designed for accurate reasoning rather than it is designed to predict the degree of truth instead of true or false. There are four modules in Fuzzy Logic to predict the system output. These are Fuzzifier, Fuzzy Rules, Fuzzy Inference Systems, and De-fuzzifications.

- Fuzzifier: Initially, crisp inputs are given to the Fuzzifier. These crisp values are actual values taken from the sensors reading.
- Fuzzy Rules: It is a set of rules like if-then-else conditions and used for decision making.
- Fuzzy Inference System: By taking fuzzy inputs and rules from the rule base, the fuzzy output ID is produced by fuzzy inference systems.
- De-fuzzification: It takes the input data from a fuzzy inference system, processes it, and produces the fuzzy output value.

In [106], the author has discussed a mobile base station using fuzzy logic. The routing with a dynamic base station is more complex compared to the static base station. In this paper, the lifetime of the network is enhanced compared to the static base station using a fuzzy controller. In [107], stable election protocol (SEP) is discussed based upon the FL control. The FL control optimizes the energy consumption required for node mobility, CH selection, and load balancing. The SEP protocol also uses weighted probability to select the cluster head and to form the cluster. A balanced energy consumption routing (BECR) is developed in [108] to enhance the network's lifetime. The author has used fuzzy C means clustering to partition the nodes into the clusters, and the node, which is the center of the cluster acts as a CH. When the energy of a cluster head drops down, the fuzzy logic system selects the other node as cluster head on a rotation basis. The protocol in [109] called a CHEF protocol that selects CHs by considering two parameters, such as energy and proximity of distance. These parameters are taken as two fuzzy parameters and the node with higher energy and locally optimal node as elected as CH. Simulation results prove the efficiency of CHEF better than LEACH by 22.7%.

The author in [110] focuses on three fuzzy metrics for CH selection that lead to enhancing network lifetime. These metrics are energy, concentration, and centrality. Still, the main drawbacks of this protocol are the lack of GPS receivers associated with the sensor nodes. There is an improvement over CHEF called F-MCHEL [111] that applies fuzzy rules based on energy and proximity of distance. A Master Cluster Head (MCH) is selected based on the highest residual energy among the CHs, gathers the data, and sends it to the base station. It is shown that F-MCHEL performs better as compared to LEACH and CHEF and brings network stability. In [117], an FL-based routing protocol is proposed that emphasizes on supercluster head (SCH) selection based on distance, energy, and centrality of CH and proves its energy efficiency. The research in [118] focuses on Type-2 Fuzzy Logic and extends the network lifetime. Many more protocols are discussed in the literature based on fuzzy logic [112–116], and we limit the discussion due to space constraints (see Tables 5–8).

5. Conclusion and future work

It is a proven fact that Wireless Sensor Networks are different from traditional networks in various aspects that demand the development of suitable communication protocols, localization techniques, data aggregation methods, scheduling mechanisms, security, fault detection, and data integrity. Machine Learning techniques help to enhance the ability of WSNs to adopt the dynamics of the environment. Furthermore,

Table 8
Summary of Fuzzy Logic-based Routing in WSNs.

Studies	Authors	Criteria	Network type	Fuzzy Parameters	Simulation Tool	Performance metrics
[106]	Abhijeet Alkesh et al.	Moving BS strategy using fuzzy logic	WSN with Mobile BS	Fuzzy controller based on energy and distance	MATLAB	No. of active nodes w.r.t rounds
[107]	Mayank Mani, and Ajay K Sharma	CH selection and node mobility	WSN with Mobile BS	Fuzzy controller based on Energy level, node density, and proximity to BS.	MATLAB	No. of alive nodes, PDR, and Remaining energy
[108]	Xin Zhao et al.	Clustering and CH selection	Homogenous WSN with static nodes	Fuzzy logic based on energy, distance from a node to the center of the cluster, and density of node itself	MATLAB	First node dies, and Average energy consumption
[109]	J.-M. Kim et al.	Selection of CH	Homogenous WSN with static nodes	Fuzzy if-Then rule based on energy, concentration, and centrality	MATLAB	No. of alive nodes and No. of clusters
[110]	I. Gupta et al.	CH selection	Homogenous WSN with static nodes	Fuzzy controller based on energy, concentration, and centrality	NRC fuzzy Java Expert System Shell (JESS)	First node dies
[111]	Tripti Sharma, and Brijesh Kumar	Master CH over CHs	Homogenous WSN with static nodes	Fuzzy inference system based on energy and proximity distance	MATLAB	No. of alive nodes, Energy, and No. of packets sent to BS
[112]	Vibha Nehra et al..	Leader selection in PEGASIS protocol	Nodes are static and have power control capabilities	A fuzzy system based on residual energy, and proximity to BS	MATLAB	No. of alive nodes, Energy and No. of packets sent to BS
[113]	Ge Ran et al..	CH selection	Homogenous WSN with static nodes	Mamdani's inference system based on distance, node density, and battery level	MATLAB	Energy consumption w.r.t rounds
[114]	Hironori Ando et al..	CH selection	Homogenous WSN with static nodes	Fuzzy controller based on number of neighbors, energy, and cluster centroid	ns-2 and MATLAB	No. of alive nodes w.r.t time
[115]	Zohre Arabi	Selection of CH and algorithm (EF-Tree, SID)	Homogenous WSN with static nodes	The fuzzy system based on energy and event	MATLAB	No. of alive nodes, PDR, Energy consumption, and Network traffic
[116]	Hoda Taheri et al.	Clustering	Homogenous WSN with static nodes	The fuzzy logic system based on node degree and centrality	MATLAB	Energy consumption, Network lifetime, and Number of CH elections
[117]	Padmalaya Nayak, D. Anurag	Super CH over CHs	WSN with static nodes and with mobile BS	The fuzzy system based on energy, mobility, and centrality	ns-2	First node dies, Half node dies, Last node dies, and end to end delay
[118]	Padmalaya Nayak, V. Bhavani	Clustering	Homogenous WSN in which all nodes are static	Type-2 fuzzy logic based on energy, distance to BS, and concentration	NetSim Simulator	First node dies, No of packets sent to BS, Energy consumption, and Throughput
[143]	Zeynab Mottaghinia and Ali Ghaffari	Enhances PDR and Reduces data transmission overhead	Mobile WSN	Fuzzy system based on distance and energy	NA	Data delivery rate and Delay

battery power is the primary source of energy in WSNs and network lifetime depends on the energy consumption of an individual sensor node. This paper gives an overview, how various machine learning techniques are used in Wireless Sensor Networks to handle the typical challenges of WSNs to enhance the network lifetime while conserving significantly less amount of energy. Moreover, the focus has been given on machine learning techniques exclusively used in Wireless Sensor Network Routing. Keeping all these elements in mind, we have extracted a few major routing protocols from the ocean of the database present in the current literature. We are hopeful that it will provide some guidelines to the current researchers to carry out their research work in this emerging area. There are several issues still open and require further investigation in WSNs. Based on the review, many more new routing protocols can be proposed by incorporating the features of distributed machine learning techniques that can make the sensor network energy aware, delay aware, and flexible enough to handle the Sensor Network constraints and dynamics of environments. Few enthusiastic points are discussed here for future research.

- *Quality of Service (QoS):* QoS is another aspect of WSNs that varies from one application to another depending on the sensor type, data type, data rate, traffic handing capability, and many more issues. In routing, addressing the QoS issues is quite a challenging task. Many real-time applications require timely delivery of data by tolerating the delay and latency. Satisfying the delay constraints, bandwidth constraints, and incorporating machine learning techniques to develop a routing protocol is quite exciting research.
- *Mobile Sink*: Mobile sink is another aspect of WSN routing research. In large scale sensor networks, data is transmitted in a multi-hop manner to reach the destination. The node closer to the sink node gets depleted soon that is known as the energy hole problem. To avoid this problem, a mobile sink is introduced to collect the data from each sensor node. But unfortunately, the mobile sink can not visit each sensor node in an extensive sensor network. So, scheduling the mobile sink in a delay-aware manner and introducing multiple mobile sinks to cover an extensive network can be done by employing ML techniques.
- *Multipath Routing:* Most of the exciting research shows that data is transmitted from a single source node to a single destination node and sensor nodes are static. When multiple source nodes and multiple destination nodes are involved in sending and receiving the data, the packet collision will occur definitely. ML techniques can be employed to handle this error.
- *Dynamic topology:* Recent studies present most of the sensor nodes are static. The inclusion of mobility introduces their position changes in WSNs. Identifying the accurate position and handling the topology dynamics, there is a requirement to develop new protocols. There is no precise mechanism to select a particular ML algorithm that is suitable for a specific application. It is the designer's responsibility to understand both the network structure and application and select appropriate ML techniques.
- *Hybrid ML Techniques:* Although many ML techniques are addressed in this paper to solve the routing issues in WSNs [1–146], many ML techniques can be combined and applied to solve the routing issues in WSNs, that yet to be explored.

Any new routing protocol can be experimented, and simulated through any type of open-source network simulators like ns-2, ns-3 simulator, and Java-based simulator. Apart from this, many professional simulators like NetSim simulator, Qualnet simulator, OMNET++ simulator, etc., can be used for practical experimentation and validation. Moreover, MATLAB can also be used as the simulation tool, and many running codes of existing routing protocols are available in Github. One may refer [117] to the experimental set up of the ns-2 simulator and in-depth analysis of a routing protocol in WSNs. A communication protocol can be measured or analyzed through various network performance parameters like average energy consumption, end to end delay, packet reception ratio, throughput, link characterization, network lifetime, etc. It can be compared with the existing protocols to prove the superiority of the protocol.

CRediT authorship contribution statement

Padmalaya Nayak: Conceptualization. **G.K. Swetha:** Formal analysis, Investigation. **Surbhi Gupta:** . **K. Madhavi:** Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work is carried on under the partial funding of JNTUH/TEQIP-III/2019/CRS/15. We are grateful to JNTUH, Hyderabad to provide us the opportunity to carry out the research work at GRIET, Hyderabad, India. We also pay our sincere thanks to the Editor and Esteemed Reviewers for their valuable suggestions and comments to enhance the quality of the survey paper.

References

[1] Jingjing Yan, Mengchu Zhou, Zhijun Ding, Recent advances in energy-efficient routing protocols for wireless sensor network: A review, IEEE Access 4 (2016) 5673–5686.
[2] Y.u. Jin Wang, X.Y. Gao, Feng Li, Hye-Jin Kim, An enhanced PEGASIS algorithm with mobile sink support for wireless sensor networks, Wirel. Commun. Mobile Comput. (2018) 1–9.
[3] R. Husain, R. Vohra, A survey on machine learning in wireless sensor networks, Int. Educ. Res. J. 3 (1) (2017).
[4] Z.A. Khan, A. Samad, A study of machine learning in wireless sensor network, Int. J. Comput. Netw. Appl. (IJCNA) 4 (4) (2017) 105–112.
[5] Y.S. Abu-Mostafa, M. Magdon-Ismail, H.-T. Lin, learning from data, AMLBook (2012).
[6] J. Qin, W. Fu, H. Gao, W.X. Zheng, Distributed *k* -means algorithm and fuzzy *c* -means algorithm for sensor networks based on multiagent consensus theory, IEEE Trans. Cybern. 47 (3) (2017) 772–783.
[7] R. El Mezouary, A. Choukri, A. Kobbane, M. El Koutbi, An energy-aware clustering approach based on the K-means method for wireless sensor networks, in: Advances in Ubiquitous Networking, Springer, 2016, pp. 325–337.
[8] A. Ray, D. De, Energy efficient clustering protocol based on k-means (EECP- K-means)-midpoint algorithm for enhanced network lifetime in wireless sensor network, IET Wirel. Sens. Syst. 6 (6) (2016) 181–191.
[9] B. Jain, G. Brar, J. Malhotra, EKMT-k-means clustering algorithmic solution for low energy consumption for wireless sensor networks based on minimum mean distance from base station, in: Networking Communication and Data Knowledge Engineering, Springer, 2018, pp. 113–123.
[10] P. Guo, J. Cao, X. Liu, Lossless in-network processing in WSNs for domain-specific monitoring applications, IEEE Trans. Ind. Inf. 13 (5) (2017) 2130–2139.
[11] X. Xu, R. Ansari, A. Khokhar, A.V. Vasilakos, Hierarchical data aggregation using compressive sensing (HDACS) in WSNs, ACM Trans. Sensor Netw. (TOSN) 11 (3) (2015) 45.
[12] P. Neamatollahi, S. Abrishami, M. Naghibzadeh, M.H.Y. Moghaddam, O. Younis, Hierarchical clustering-task scheduling policy in cluster-based wireless sensor net- works, IEEE Trans. Ind. Inf. 14 (5) (2018) 1876–1886.
[13] R. Zhang, J. Pan, D. Xie, F. Wang, NDCMC: A hybrid data collection approach for large-scale WSNs using mobile element and hierarchical clustering, IEEE Internet Things J. 3 (4) (2016) 533–543.
[14] R. Zhang, J. Pan, J. Liu, D. Xie, A hybrid approach using mobile element and hierarchical clustering for data collection in WSNs, in: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), 2015, pp. 1566–1571.
[15] Sadia Waheed Awan, Sajid Saleem, Hierarchical Clustering algorithms for heterogeneous energy harvesting wireless sensor networks, in: 2016 International Symposium on Wireless Communication Systems (ISWCS).
[16] X. Li, S. Ding, Y. Li, Outlier suppression via non-convex robust PCA for efficient localization in wireless sensor networks, IEEE Sens. J. 17 (21) (2017) 7053–7063.
[17] A. Morell, A. Correa, M. Barceló, J.L. Vicario, Data aggregation and principal com- ponent analysis in WSNs, IEEE Trans. Wirel. Commun. 15 (6) (2016) 3908–3919.
[18] T. Yu, X. Wang, A. Shami, Recursive principal component analysis based data outlier detection and sensor data aggregation in IoT systems, IEEE Internet Things J. 4 (6) (2017) 2207–2216.
[19] M. Wu, L. Tan, N. Xiong, Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications, Inf. Sci. 329 (2016) 800–818.
[20] P. Oikonomou, A. Botsialas, A. Olziersky, I. Kazas, I. Stratakos, S. Katsikas, D. Dimas, K. Mermikli, G. Sotiropoulos, D. Goustouridis, et al., A wireless sensing system for monitoring the workplace environment of an industrial installation, Sens. Actuators B 224 (2016) 266–274.
[21] M.R. Islam, J. Uddin, J.-M. Kim, Acoustic emission sensor network based fault diagnosis of induction motors using a gabor filter and multiclass support vector machines, Adhoc. Sens. Wirel. Netw. 34 (2016) 273–287.
[22] Q.-Y. Sun, Y.-M. Sun, X.-J. Liu, Y.-X. Xie, X.-G. Chen, Study on fault diagnosis algo- rithm in WSN nodes based on RPCA model and SVDD for multi-class classification, Cluster Comput. (2018) 1–15.
[23] G. Gennarelli, F. Soldovieri, Performance analysis of incoherent RF tomography using wireless sensor networks, IEEE Trans. Geosci. Remote Sens. 54 (5) (2016) 2722–2732.
[24] W. Peizhuang, Pattern recognition with fuzzy objective function algorithms (james c. bezdek), SIAM Rev. 25 (3) (1983) 442.
[25] F. Zhu, J. Wei, Localization algorithm for large-scale wireless sensor networks based on FCMTSR–support vector machine, Int. J. Distrib. Sens. Netw. 12 (10) (2016) 1–12.
[26] M. Bernas, B. P ł aczek, Fully connected neural networks ensemble with signal strength clustering for indoor localization in wireless sensor networks, Int. J. Distrib. Sens. Netw. 11 (12) (2015) 1–10.
[27] Padmalaya Nayak, Praneeth Reddy, A bio-inspired routing protocol for wireless sensor network to minimize the energy consumption, IET Wirel. Sens. Syst. 10 (5) (2020).
[28] J. Qin, W. Fu, H. Gao, W.X. Zheng, Distributed *k*-means algorithm and fuzzy *c* -means algorithm for sensor networks based on multiagent consensus theory, IEEE Trans. Cybern. 47 (3) (2017) 772–783.
[29] A. Mehmood, Z. Lv, J. Lloret, M.M. Umar, ELDC: An artificial neural network based energy-efficient and robust routing scheme for pollution monitoring in WSNs, IEEE Trans. Emerg. Top Comput. PP (99) (2017) 1–8.
[30] M.S. Gharajeh, S. Khanmohammadi, DFRTP: Dynamic 3D fuzzy routing based on traffic probability in wireless sensor networks, IET Wirel. Sens. Syst. 6 (6) (2016) 211–219.
[31] J.R. Srivastava, T. Sudarshan, A genetic fuzzy system based optimized zone based energy efficient routing protocol for mobile sensor networks (OZEEP), Appl. Soft Comput. 37 (2015) 863–886.
[32] Y. Lee, Classification of node degree based on deep learning and routing method applied for virtual route assignment, Ad Hoc Netw. 58 (2017) 70–85.
[33] F. Khan, S. Memon, S.H. Jokhio, Support vector machine-based energy aware rout- ing in wireless sensor networks, in: 2016 2nd International Conference on Robotics and Artificial Intelligence (ICRAI), 2016, pp. 1–4.
[34] V. Jafarizadeh, A. Keshavarzi, T. Derikvand, Efficient cluster head selection using naïve bayes classifier for wireless sensor networks, Wirel. Netw. 23 (3) (2017) 779–785.
[35] Z. Liu, M. Zhang, J. Cui, An adaptive data collection algorithm based on a bayesian compressed sensing framework, Sensors 14 (5) (2014) 8330–8349.
[36] F. Kazemeyni, O. Owe, E.B. Johnsen, I. Balasingham, Formal modeling and analysis of learning-based routing in mobile wireless sensor networks, in: Integration of Reusable Systems, Springer, 2014, pp. 127–150.
[37] X. Song, C. Wang, J. Gao, X. Hu, DLRDG: distributed linear regression-based hierarchical data gathering framework in wireless sensor network, Neural Comput. Appl. 23 (7–8) (2013) 1999–2013.
[38] L. Gispan, A. Leshem, Y. Be`ery, Decentralized estimation of regression coefficients in sensor networks, Digit Signal Process. 68 (2017) 16–23.
[39] C. Habib, A. Makhoul, R. Darazi, C. Salim, Self-adaptive data collection and fusion for health monitoring based on body sensor networks, IEEE Trans. Ind. Inf. 12 (6) (2016) 2342–2352.
[40] Y. Li, L.E. Parker, Nearest neighbor imputation using spatial–temporal correlations in wireless sensor networks, Inf. Fusion 15 (2014) 64–79.
[41] A. De Paola, P. Ferraro, S. Gaglio, G.L. Re, S.K. Das, An adaptive bayesian system for context-aware data fusion in smart environments, IEEE Trans. Mob. Comput. 16 (6) (2017) 1502–1515.
[42] C. Wang, E. Bertino, Sensor network provenance compression using dynamic bayesian networks, ACM Trans. Sensor Netw. (TOSN) 13 (1) (2017) 5.1–5.32.
[43] F. Edwards-Murphy, M. Magno, P.M. Whelan, J. OHalloran, E.M. Popovici, b+ WSN: smart beehive with preliminary decision tree analysis for agriculture and honey bee health monitoring, Comput. Electron. Agric. 124 (2016) 211–219.

[44] W. Zhao, S. Su, F. Shao, Improved DV-hop algorithm using locally weighted linear regression in anisotropic wireless sensor networks, Wirel. Personal Commun. 98 (4) (2018) 3335–3353.
[45] S.S. Banihashemian, F. Adibnia, M.A. Sarram, A new range-free and storage-efficient localization algorithm using neural networks in wireless sensor networks, Wirel. Personal Commun. 98 (1) (2018) 1547–1568.
[46] S.K. Gharghan, R. Nordin, M. Ismail, J.A. Ali, Accurate wireless sensor localization technique based on hybrid PSO-ANN algorithm for indoor and outdoor track cycling, IEEE Sens. J. 16 (2) (2016) 529–541.
[47] J. Kang, Y.J. Park, J. Lee, S.H. Wang, D.S. Eom, Novel leakage detection by ensemble CNN-SVM and graph-based localization in water distribution systems, IEEE Trans. Ind. Electron. 65 (5) (2018) 4279–4289.
[48] T. Tang, H. Liu, H. Song, B. Peng, Support vector machine based range-free localization algorithm in wireless sensor network, in: International Conference on Machine Learning and Intelligent Communications, 2016, pp. 150–158.
[49] T.L.T. Nguyen, F. Septier, H. Rajaona, G.W. Peters, I. Nevat, Y. Delignon, A bayesian perspective on multiple source localization in wireless sensor networks, IEEE Trans. Signal Process. 64 (7) (2016) 1684–1699.
[50] Z. Wang, H. Liu, S. Xu, X. Bu, J. An, Bayesian device-free localization and tracking in a binary RF sensor network, Sensors 17 (5) (2017) 1–21.
[51] Z. Xiahou, X. Zhang, Adaptive localization in wireless sensor network through bayesian compressive sensing, Int. J. Distrib. Sens. Netw. 2015 (2015) 1–15.
[52] J. Yoo, W. Kim, H.J. Kim, Distributed estimation using online semi-supervised particle filter for mobile sensor networks, IET Control Theory Appl. 9 (3) (2015) 418–427.
[53] B. Yang, J. Xu, J. Yang, M. Li, Localization algorithm in wireless sensor networks based on semi-supervised manifold learning and its application, Cluster Comput. 13 (4) (2010) 435–446.
[54] Michele Chincoli, Antonio Liotta, Self-learning power control in wireless sensor networks, MDPI Sens. (2018).
[55] P. Braca, P. Willett, K.D. LePage, S. Marano, V. Matta, Bayesian tracking in underwater wireless sensor networks with port-starboard ambiguity, IEEE Trans. Signal Process. 62 (7) (2014) 1864–1878.
[56] V.P. Illiano, E.C. Lupu, Detecting malicious data injections in event detection wire- less sensor networks, IEEE Trans. Netw. Serv. Manage. 12 (3) (2015) 496–510.
[57] W. Sun, X. Yuan, J. Wang, Q. Li, L. Chen, D. Mu, End-to-end data delivery reliability model for estimating and optimizing the link quality of industrial WSNs, IEEE Trans. Autom. Sci. Eng. 15 (3) (2018) 1127–1137.
[58] X. Chang, J. Huang, S. Liu, G. Xing, H. Zhang, J. Wang, L. Huang, Y. Zhuang, Accu- racy-aware interference modeling and measurement in wireless sensor networks, IEEE Trans. Mob. Comput. 15 (2) (2016) 278–291.
[59] M.A. Alsheikh, S. Lin, D. Niyato, H.P. Tan, Rate-distortion balanced data compression for wireless sensor networks, IEEE Sens. J. 16 (12) (2016) 5072–5083.
[60] A.A. Rezaee, F. Pasandideh, A fuzzy congestion control protocol based on active queue management in wireless sensor networks with medical applications, Wirel. Personal Commun. 98 (1) (2018) 815–842.
[61] M.R. Islam, J. Uddin, J.-M. Kim, Acoustic emission sensor network based fault diagnosis of induction motors using a gabor filter and multiclass support vector ma- chines, Adhoc Sens. Wirel. Netw. 34 (2016) 273–287.
[62] M. Gholipour, A.T. Haghighat, M.R. Meybodi, Hop-by-Hop congestion avoidance in wireless sensor networks based on genetic support vector machine, Neurocomputing 223 (2017) 63–76.
[63] W. Kim, M.S. Stankovi, K.H. Johansson, H.J. Kim, A distributed support vector machine learning over wireless sensor networks, IEEE Trans. Cybern 45 (11) (2015) 2599–2611.
[64] M. Xie, J. Hu, S. Han, H.-H. Chen, Scalable hypergrid k-NN-based online anomaly detection in wireless sensor networks, IEEE Trans. Parallel Distrib. Syst. 24 (8) (2013) 1661–1670.
[65] B. Yang, Y. Lei, B. Yan, Distributed multi-human location algorithm using naive bayes classifier for a binary pyroelectric infrared sensor tracking system, IEEE Sens. J. 16 (1) (2016) 216–223.
[66] J. Shu, S. Liu, L. Liu, L. Zhan, G. Hu, Research on link quality estimation mechanism for wireless sensor networks based on support vector machine, Chin. J. Electr. 26 (2) (2017) 377–384.
[67] B. Alotaibi, K. Elleithy, A new MAC address spoofing detection technique based on random forests, Sensors 16 (3) (2016) 1–14.
[68] W. Elghazel, K. Medjaher, N. Zerhouni, J. Bahi, A. Farhat, C. Guyeux, M. Hakem, Random forests for industrial device functioning diagnostics using wireless sensor networks, in: Aerospace Conference, 2015 IEEE, IEEE, 2015, pp. 1–9.
[69] Y. Li, H. Chen, M. Lv, Y. Li, Event-based k-nearest neighbors query processing over distributed sensory data using fuzzy sets, Soft Comput. (2017) 1–13.
[70] Y. Wang, A. Yang, X. Chen, P. Wang, Y. Wang, H. Yang, A deep learning approach for blind drift calibration of sensor networks, IEEE Sens. J. 17 (13) (2017) 4158–4171.
[71] M.M. Hassan, M.Z. Uddin, A. Mohamed, A. Almogren, A robust human activity recognition system using smartphone sensors and deep learning, Future Gener. Comput. Syst. 81 (2018) 307–313.
[72] F. Chen, Z. Fu, Z. Yang, Wind power generation fault diagnosis based on deep learning model in internet of things (IoT) with clusters, Cluster Comput. (2018) 1–13.
[73] F.K. Shaikh, S. Zeadally, Energy harvesting in wireless sensor networks: a comprehensive review, Renewable Sustainable Energy Rev. 55 (2016) 1041–1054.
[74] S. Kosunalp, A new energy prediction algorithm for energy-harvesting wireless sen- sor networks with Q-learning, IEEE Access 4 (2016) 5755–5763.
[75] J. Wang, Y. Cao, B. Li, H.-J. Kim, S. Lee, Particle swarm optimization based cluster- ing algorithm with mobile sink for WSNs, Future Gener. Comput. Syst. 76 (2017) 452–457.
[76] K.D. Praveen, T. Amgoth, C.S.R. Annavarapu, ACO-based mobile sink path determination for wireless sensor networks under non-uniform data constraints, Appl. Soft Comput. 69 (2018) 528–540.
[77] G. Yogarajan, T. Revathi, Nature inspired discrete firefly algorithm for optimal mobile data gathering in wireless sensor networks, Wirel. Networks (2017) 1–15.
[78] S.-H. Moon, S. Park, S.-J. Han, Energy efficient data collection in sink-centric wireless sensor networks: a cluster-ring approach, Comput. Commun. 101 (2017) 12–25.
[79] A.P. Renold, S. Chandrakala, MRL-SCSO: Multi-agent reinforcement learning-based self-configuration and self-optimization protocol for unattended wireless sensor networks, Wirel. Personal Commun. 96 (4) (2017) 5061–5079.
[80] R. Kirthana, V.V. Bhargavi, Online incremental learning algorithm for anomaly detection and prediction in health care, In: International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, 2014.
[81] S.A. Haque, M. Rahman, S.M. Aziz, Sensor anomaly detection in wireless sensor networks for healthcare, Sensors 15 (4) (2015) 8764–8786.
[82] Z. Feng, J. Fu, D. Du, F. Li, S. Sun, A new approach of anomaly detection in wireless sensor networks using support vector data description, Int. J. Distrib. Sens. Netw. 13 (1) (2017) 1–14.
[83] N. Shahid, I.H. Naqvi, S.B. Qaisar, One-class support vector machines: analysis of outlier detection for wireless sensor networks in harsh environments, Artif. Intell. Rev. 43 (4) (2015) 515–563.
[84] A. Garofalo, C. Di Sarno, V. Formicola, Enhancing intrusion detection in wireless sensor networks through decision trees, Depend. Comput. (2013) 1–15.
[85] Yong Wang, Margaret Martonosi, Li-Shiuan Peh, A supervised learning approach for routing optimizations in wireless sensor networks, in: Proceedings of the 2nd International Workshop on Multi-hop ad hoc Networks: From Theory to Reality, ACM, 2006.
[86] Yong Wang, Margaret Martonosi, Li-Shiuan Peh, Predicting link quality using supervised learning in wireless sensor networks, ACM SIGMOBILE Mobile Comput. Commun. Rev. 11 (3) (2007) 71–83.
[87] Jeroen L. de Jung, Wilbert van Norden, Fok Bolderheij, Leon Rothkrantz, Intelligent task scheduling in sensor networks. Proceedings of IEEE 2005 8th International Conference on Information Fusion, 2005.
[88] Mao Ching Foo, Hock Beng Lim, Yulian Zeng, Vinh The Lam, R. Teo, Gee Wah Ng, Impact of distributed resource allocation in sensor networks, in: 2005 Intelligent Sensors, Sensor Networks and Information Processing Conference (ISSNIP), Melbourne, Australia, 2005, pp. 69–74.
[89] H.B. Lim, Vinh The. Lam, Mao Ching Foo, Yulian Zeng, An adaptive distributed resource allocation scheme for sensor networks, MSN 2006, LNCS 4325, 2006, pp. 770-781.
[90] A. Mehmood, Z. Lv, J. Lloret, M.M. Umar, ELDC: An artificial neural network based energy-efficient and robust routing scheme for pollution monitoring in WSNs, IEEE Trans. Emerg. Top Comput. PP (99) (2017) 1–8.
[91] Attiah, Afraa, et al., An evolutionary game for efficient routing in wireless sensor networks, in: Global Communications Conference (GLOBECOM), 2016 IEEE. IEEE, 2016.
[92] Ketema Adere Gemeda, Gabriele Gianini, Mulugeta Libsie, An evolutionary cluster-game approach for Wireless Sensor Networks innoncollaborative settings, Pervasive Mobile Comput. 42 (2017) 209–225.
[93] Zhou Ruyan, Chen Ming, Feng Guofu, Liu Huifang, He Shijun, Genetic clustering route algorithm in WSN, in: IEEE 2010 Sixth International Conference on Natural Computation (ICNC 2010).
[94] Wu Yin, Wenbo Liu, Routing protocol based on genetic algorithm for energy harvesting wireless sensor networks, IEEE IET Wirel. Sens. Syst. 3 (2) (2013) 112–118.
[95] Omar Banimelhem, Moad Mowafi, Eyad Taqieddin, Fahed Awad, Manar Al Rawabdeh, An efficient clustering approach using genetic algorithm and node mobility in wireless sensor networks, in: IEEE 2014 11th International Symposium on Wireless Communications Systems (ISWCS).
[96] Ataul Bari, Shamsul Wazed, Arunita Jaekel, Subir Bandyopadhyay, A genetic algorithm-based approach for energy-efficient routing in two-tiered sensor networks, Ad Hoc Networks 7 (2009) 665–676.
[97] N. Srinivas, K. Deb, Multiobjective optimization using nondominated sorting in genetic algorithms, J. Evolut. Comput. 2 (3) (1995) 221–248.
[98] Y. Pan, X. Liu, Energy-efficient lifetime maximization and sleeping scheduling supporting data fusion and QoS in Multi-Sensor Net, Signal Process. 87 (12) (2007) 2949–2964.
[99] Q. Qiu, Q. Wu, D. Burns, D. Holzhauer, Lifetime aware resource management for sensor network using distributed genetic algorithm, in: ISLPED'06, ACM Press, New York, NY, 2006, pp. 191–196.
[100] Navrati Saxena, Abhishek Roy, Jitae Shin, QuESt: a QoS-based energy-efficient sensor Routing Protocol, Wirel. Commun. Mobile Comput. J. 9 (3) (2009) 417–426.
[101] Jin M. Zhou, A.S. Wu, Sensor network optimization using a genetic algorithm, in: Proceedings of 7th World Multiconference on Systemics, Cybernetics and Informatics, 2003.
[102] P. Nayak, V. Bhabani, A genetic algorithm based clustering algorithm for wireless sensor network to enhance the network lifetime, in: IEEE CONFLUENCE 2017, Amity University, Delhi, 12–13 Jan 2017.
[103] Veena Trivedi, Padmalaya Nayak, Modified AODV Using Genetic algorithm, Algorithm to minimize energy consumption in MANET, IJITEE 8 (7s2) (2019) 525–530.

[104] P. Kuila, P.K. Jana, Energy-efficient clustering and routing algorithms for wireless sensor networks: particle swarm optimization approach, Eng. Appl. Artif. Intell. 33 (2014) 127–140.
[105] Y. Sun, W. Dong, Y. Chen, An improved routing algorithm based on ant colony optimization in wireless sensor networks, IEEE Commun. Lett. 21 (6) (2017) 1317–1320.
[106] Abhijeet Alkesh, Ashutosh Kumar Singh, and N. Purohit, A moving base station strategy using fuzzy logic for lifetime enhancement in wireless sensor network, in: 2011 International Conference on Communication Systems and Network Technologies.
[107] Mayank Mani, Ajay K. Sharma, Modified approach for routing and clustering in Sensor Network using Fuzzy Logic Control, in: IEEE 2013 Sixth International Conference on Contemporary Computing (IC3), Noida, India, pp. 102–107.
[108] Xin Zhao, Zhiqiang Wei, Yanping Cong, Bo Yin, A balances energy consumption clustering routing protocol for a wireless sensor network. 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), 2018.
[109] J.-M. Kim, S.-H. Park, Y.-J. Han, T.-M. Chung, CHEF: Cluster head election mechanism using fuzzy logic in wireless sensor networks, in: Proc. 10th ICACT, Feb. 2008, 2008, pp. 654–659.
[110] I. Gupta, D. Riordan, S. Sampalli, Cluster-head election using fuzzy logic for wireless sensor networks, in: Proc. 3rd Annu. Commun. Netw Services Res. Conf., May 2005, pp. 255–260.
[111] Tripti Sharma, Brijesh Kumar, F-MCHEL: Fuzzy based master cluster head election leach protocol in wireless sensor network, Int. J. Comput. Sci. Telecommun. 3 (10) (2012) 8–13.
[112] Vibha Nehra, Raju Pal, Ajay K. Sharma, Fuzzy-based leader selection for topology controlled PEGASIS protocol for lifetime enhancement in wireless sensor network, Int. J. Comput. Technol. 4 (3) (2013) 755–764.
[113] Ge Ran, Huazhong Zhang, Shulan Gong, Improving on LEACH protocol of wireless sensor networks using fuzzy logic, J. Inf. Comput. Sci. 7 (3) (2010) 767–775.
[114] Hironori Ando, Leonard Barolli, Arjan Durresi, Fatos Xhafa, Akio Koyama, An intelligent fuzzy-based cluster head selection system for WSNs and its performance evaluation for D3N parameter, in: 2010 International Conference on Broadband, Wireless Computing, Communication and Applications, 2010, pp. 648–653.
[115] Zohre Arabi, HERF: A hybrid energy efficient routing using a fuzzy method in wireless sensor networks, in: International Conference on Intelligent and Advanced Systems (ICIAS), 2010, pp. 1–6.
[116] Hoda Taheri, et al., An energy-aware distributed clustering protocol in wireless sensor networks using fuzzy logic, Ad hoc Networks 10 (2012) 1469–1481.
[117] Padmalaya Nayak, D. Anurag, Fuzzy logic based routing protocol for wireless sensor network to extend the network lifetime, IEEE Sens. J. 16 (1) (2016) 137–144.
[118] Padmalaya Nayak, V. Bhavani, Energy efficient clustering algorithm for multi-hop wireless sensor network using type 2 fuzzy logic, IEEE Sens. J. 17 (14) (2017) 4492–4499.
[119] Neda Enami, Reza Askari, Moghadam, Energy-based clustering self organizing map protocol for extending wireless sensor networks lifetime and coverage, Can. J. Multimedia Wirel. Networks 1 (4) (2010).
[120] Chiranjib Patra, Parama Bhaumik, Matangini Chattopadhyay, Anjan Guha Roy, Using self organizing map in wireless sensor network for designing energy-efficient topologies, in: IEEE 2011 2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE), Chennai, India.
[121] Michal Turčaník, Network Routing by Artificial Neural Network, IEEE 2012 Military Communications and Information Systems Conference (MCC), Gdansk, Poland, 8–9 Oct. 2012.
[122] Mohit Mittal, Krishan Kumar, Network lifetime enhancement of homogeneous sensor network using ART1 neural network, in: IEEE 2014 Sixth International Conference on Computational Intelligence and Communication Networks, 14–26 Nov. 2014, Bhopal, pp. 472–475.
[123] Tao Li, Feng Ruan, Zhiyong Fan, Jin Wang, Jeong-Uk Kim, An improved PEGASIS protocol for wireless sensor network. IEEE 2015 3rd International Conference on Computer and Computing Science, 2015.
[124] Junling Li, Danpu Liu, An energy-aware distributed clustering routing protocol for energy harvesting wireless sensor networks, in: 2016 IEEE/CIC International Conference on Communications in China (ICCC).
[125] Deepshikha, Priyanka Arora, Varsha, Enhanced NN based RZ LEACH using hybrid ACO/PSO based routing for WSNs, in; IEEE 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT).
[126] Mohammad Abu Alsheikh, Shaowei Lin, Dusit Niyato, Hwee-Pink Tan, Machine learning in wireless sensor networks: algorithms, strategies, and applications, IEEE Commun. Surv. Tutorials 4 (4) (2014) 1996–2018.
[127] S. Dong, P. Agrawal, K. Sivalingam, Reinforcement learning-based geographic routing protocol for UWB wireless sensor network, in: Global Telecommunications Conference, IEEE, 2007, pp. 652–656.
[128] G. Aiello, G. Rogerson, Ultra-wideband wireless systems, IEEE Microwave Magazine 4 (2) (2003) 36–47.
[129] T. Zhang, R. Ramakrishnan, M. Livny, BIRCH: An efficient data clustering method for very large databases, ACM SIGMOD Record 25 (2) (1996) 103–114.
[130] S. Guha, R. Rastogi, K. Shim, CURE: An efficient clustering algorithm for large databases, in: Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data, ser. SIGMOD '98, ACM, New York, NY, USA, 1998, pp. 73–84.
[131] K. Crammer, A. Kulesza, M. Dredze, Adaptive regularization of weight vectors, Mach. Learn. 91 (2) (2013) 155–187.
[132] L. Yang, R. Jin, J. Ye, Online learning by ellipsoid method, in: Proceedings of the 26th Annual International Conference on Machine Learning, ser. ICML '09, ACM, New York, NY, USA, 2009, pp. 1153–1160.
[133] D.D. Andrea Kulakov, Distributed data processing in wireless sensor networks based on artificial neural-networks algorithms, In: Proceedings of 10th IEEE Symphosiums on Computers and Communication (ISCC), NW Washington, DC, United States, 2005.
[134] Andrea Kulakov, Danco Davcev, Intelligent data acquisition and processing using wavelet neural-networks, in: 2005 IEEE Intelligent Data Acquisition and Advanced Computing Systems: Technology and Application, Bulgaria, 5–7 Sept 2005.
[135] Seunggye Hwang, Rong rang, Janghoon Yang, Dong Ku Kim, Multivariate Bayesian compressive sensing in wireless sensor networks, IEEE Sens. J. 16 (2016) issues 7.
[136] X. Jiang, J. Taneja, J. Ortiz, A. Tavakoli, P. Dutta, J. Jeong, D.E. Culler, P. Levis, S. Shenker, et al., An architecture for energy management in wireless sensor networks, SIGBED Rev. 4 (3) (2007) 31–36.
[137] D. Praveen Kumar, Tarachand Amgoth, Chandra Sekhara Rao Annavarapu, Machine learning algorithms for wireless sensor networks: A survey, Inf. Fusion 49 (2019) 1–25.
[138] R. Sun, S. Tatsumi, G. Zhao, Q-MAP: A novel multicast routing method in wireless ad hoc networks with multiagent reinforcement learning, in: Region 10 Conference on Computers, Communications, Control and Power Engineering, vol. 1, 2002, pp. 667–670.
[139] Yu. Jin Wang, Wei Liu Gao, Wu Webing, Se-Jung Lim, An asynchronous clustering and mobile data gathering schema based on timer mechanism in wireless sensor networks, Comput. Mater. Continua 58 (3) (2019) 711–725.
[140] Jin Wang, Yu Gao, Chang Zhou, R.S. Sherratt, Lei Wang, Optimal coverage multi-path scheduling scheme with multiple mobile sinks for WSNs, Comput. Mater. Continua 62 (2) (2020) 695–711.
[141] Melika Khabiri, Ali Ghaffari, Energy-aware clustering-based routing in wireless sensor networks using cuckoo optimization algorithm, Wirel. Pers. Commun. 98 (3) (2018) 2473–2495.
[142] Shamineh Tabibi, Ali Ghaffari, Energy-efficient routing mechanism for mobile sink in wireless sensor networks using particle swarm optimization algorithm, Wirel. Pers. Commun. 104 (1) (2019) 199–216.
[143] Zeynab Mottaghinia, Ali Ghaffari, Fuzzy logic based distance and energy-aware routing protocol in delay-tolerant mobile sensor networks, Wirel. Pers. Commun. 100 (3) (2018) 957–976.
[144] Jing Bi, Haitao Yuan, Mengchu Zhou, Temporal prediction of multiapplication consolidated workloads in distributed clouds, IEEE Trans. Autom. Sci. Eng. 16 (4) (2019) 1763–1773.
[145] Jing Bi, Haitao Yuan, LiBo Zhang, Jia Zhang, SGW-SCN: An integrated machine learning approach for workload forecasting in Geo-Distributed cloud data centers, Inf. Sci. 481 (2019) 57–68.
[146] Jinghui Zhong, Zhixing Huang, Liang Feng, Du Wan, Ying Li, A hyper-heuristic framework for lifetime maximization in wireless sensor networks with a mobile sink, IEEE/CAA J. Automat. Sin. 7 (1) (2020) 223–226.