J Inf Process Syst, Vol.15, No.4, pp.833~852, August 2019
https://doi.org/10.3745/JIPS.04.0125

ISSN 1976-913X (Print)
ISSN 2092-805X (Electronic)

JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Selection of Machine Learning Techniques for Network Lifetime Parameters and Synchronization Issues in Wireless Networks

Nimmagadda Srilakshmi* and Arun Kumar Sangaiah*

Abstract

In real time applications, due to their effective cost and small size, wireless networks play an important role in receiving particular data and transmitting it to a base station for analysis, a process that can be easily deployed. Due to various internal and external factors, networks can change dynamically, which impacts the localisation of nodes, delays, routing mechanisms, geographical coverage, cross-layer design, the quality of links, fault detection, and quality of service, among others. Conventional methods were programmed, for static networks which made it difficult for networks to respond dynamically. Here, machine learning strategies can be applied for dynamic networks effecting self-learning and developing tools to react quickly and efficiently, with less human intervention and reprogramming. In this paper, we present a wireless networks survey based on different machine learning algorithms and network lifetime parameters, and include the advantages and drawbacks of such a system. Furthermore, we present learning algorithms and techniques for congestion, synchronisation, energy harvesting, and for scheduling mobile sinks. Finally, we present a statistical evaluation of the survey, the motive for choosing specific techniques to deal with wireless network problems, and a brief discussion on the challenges inherent in this area of research.

Keywords

Congestion, Energy Harvesting, Machine Learning Algorithms, Network Lifetime, Wireless Networks

1. Introduction

In some applications, a large number of intermediate nodes are present, and managing them requires efficient and scalable algorithms. Due to external factors, networks can change dynamically, which causes the localisation of nodes, delays, routing mechanisms, geographical coverage, cross layer designing, quality of link, fault detection, and quality of service, among others. Applying machine learning according to recent advances in technology has solved many challenges to improve network performance, and requiring less human intervention and programming. Recent advancements in technology and applications are important for effecting integration between physical systems, the Internet of Things, machine to machine cooperation, cloud to meet the requirements of user.

The following are examples of the application of machine learning. (1) Deciding the optimum number

Manuscript received February 1, 2019; first revision April 9, 2019; accepted April 18, 2019.
Corresponding Author: Arun Kumar Sangaiah (sarunkumar@vit.ac.in)
* School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India ({nimmagadda.srilakshmi2015, sarunkumar}@vit.ac.in)

of nodes required to overcome target area coverage problems. (2) Forecasting the quantity of energy consumed within a given time period, for long-life and self-powered systems to overcome the energy harvesting problem. (3) Machine learning algorithms that can rapidly and efficiently establish accurate localisation of nodes to overcome the dynamic nature of network organisation due to external or internal factors. (4) Accumulating to the cloud gaming in multiplayer cooperative scenario to overcome the synchronising of data with respect to their cloud data. (5) Improving efficiency by overcoming the differentiating problem between the fault and normal nodes in a network. (6) Enhancing network lifetime and system performance through dynamic routing strategies to overcome the dynamic behaviour of networks. (7) The dimensionality of data can be lowered to overcome overheads incurred by transmitting complete data at a base station or to the cluster heads in a cluster.

We organize the survey paper as follows. In Section 1 introduction to machine learning algorithms to wireless networks and cloud is presented. In Section 2 background of different types of machine learning techniques are discussed. In Section 3 different applications and their issues in networks with machine learning techniques are surveyed. In section 4 limitations and statistical analysis are explained. In Section 5 open challenges for research in machine learning based networks and cloud are presented. Finally, the conclusions are drawn for the survey in Section 6.

2. Machine Learning Techniques

2.1 Supervised Learning

Supervised learning is an important data processing technique in the field of machine learning. When training a system, it provides a set of inputs and outputs, i.e., a data set with labels, and establishes a link between them. This learning algorithm provides dependency links and relationships among inputs, and foresees outputs. Table 1 shows comparison of various machine learning techniques with different parameters.

2.2 Unsupervised Learning

The unsupervised learning technique is used for classifying data into similar patterns, reducing the size of data, forming clusters, and anomaly detection. It is associated with given inputs and as such, it involves no unlabelled output. This approach solves challenges related to connectivity problems, routing, data aggregation, and anomaly detection. It is divided into dimensionality reduction, such as singular value decomposition, independent component analysis, principle component analysis, and clustering, e.g., fuzzy c-means, k-means, and hierarchical. Table 2 depicts comparisons of various clustering algorithms.

2.3 Semi-Supervised Learning

Semi-supervised learning works on both supervised (labelled) and unsupervised (unlabelled) data. In real world applications based on semi-supervised learning, classification is performed partially on labelled data, and regression on unlabelled data. The important factor is to predict whether data is labelled or unlabelled in training datasets and future datasets. In video surveillance, speech recognition, classifying web content, natural language processing, protein sequence classification, and spam filtering applications, and to solve fault detection and localisation in wireless networks, this learning technique is employed.

Table 1. Comparison of various machine learning techniques with different parameters

Specification type	Decision tree	Reinforcement learning	ANN	Deep learning	SVM	Bayesian	K-NN
Parameter handling	Very good	Very good	Poor	Good	Poor	Best	Very good
Speed of learning	Very good	Good	Poor	Poor	Poor	Best	Best
Accuracy	Good	Good	Very good	Very good	Best	Poor	Good
Speed of classification	Best	Best	Best	Best	Best	Best	Poor
Missing values handling	Very good	Good	Best	Good	Good	Best	Good
Redundant variables handling	Good	Good	Good	Good	Very good	Poor	Good
Noise handling	Good	Very good	Good	Very good	Good	Very good	Poor
Independent variables handling	Good	Good	Very good	Very good	Very good	Poor	Poor
Irrelevant variables handling	Very good	Very good	Poor	Good	Best	Good	Good
Dealing over fitting	Good	Good	Poor	Poor	Good	Very good	Very good

Table 2. Comparisons of various clustering algorithms

Specification type	K-means clustering	Fuzzy c-means clustering	Hierarchical clustering
Clustering speed	Fast	Slow	Fast
Accuracy	Low	High	High
Accuracy of prediction	High	Low	Low
Quality	High	Moderate	Moderate
Sensitivity	High	Low	Low
Randomness in results	Moderate	Moderate	Good
Performance	High	Moderate	Moderate

2.4 Reinforcement Learning

Reinforcement learning gathers data by continuously interacting with surrounding data and performing all necessary actions by drawing conclusions thereby improving performance of system by obtaining optimal results. Q-learning is a type of reinforcement learning that generates observation sequences as state-action-rewards. Based on behaviour, e.g., assertive, aggressive, passive-aggressive, and passive, reinforcement learning can be categorised as negative, positive, extinction, and punishment types. Visualization of reinforcement learning is depicted in Fig. 1.

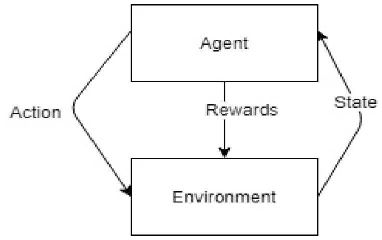

Fig. 1. Reinforcement learning visualization.

2.5 Evolutionary Computation

Evolutionary computation is an artificial intelligence sub-category that uses several optimisation techniques for computation that are inspired by biological evolution and nature, where a solution is obtained using several iterations. To obtain optimal results, evolutionary computation rejects less fitting solutions in iterations according to a trial and error process. This includes ant colony optimisation, genetic programming, genetic algorithms, evolutionary programming, evolutionary algorithms, artificial immune systems, artificial bee colonies, particle swarm optimisation, and firefly algorithms. This method is used to solve coverage, localisation, target tracking, routing, and mobile sink issues in wireless networks.

3. Wireless Networks – Machine Learning Algorithms

In this section, machine learning techniques for coping with challenges in wireless networks are discussed, as well as their advantages, alongside existing approaches depicted in their respective tabular forms.

3.1 Localization

Recognising the geographical, physical location of a wireless node manually, or by a global positioning system by sending beacon or anchor nodes, is known as localisation. This can be based on proximity, distance and angle, range, or location of nodes Continuous configuring and programming is needed for a dynamically changing network, where machine learning techniques must be applied to improve the accuracy of pinpointing location. This presents various advantages, e.g., anchor and unknown nodes can easily be found using machine learning algorithms in a network, to create clusters, and training them separately. Table 3 draws machine learning based approaches for localization in wireless networks.

Table 3. Machine learning based approaches for localization

Machine learning mechanism	Complexity	Mobility of nodes	Network	Contributions	Ref.
K-means & fuzzy c-means	High	Static	Centralized	Accuracy is improved	[1]
Principle component analysis	Moderate	Static	Distributed	Outlier detection	[2]
Regression	Moderate	Static	Distributed	Accuracy is improved	[3]
ANN	Low	Static	Distributed	Error rate is reduced	[4]
	Moderate	Static	Centralized	Accuracy is improved	[5]
Fuzzy logic	Low	Static	Distributed	Time complexity is reduced	[6]
	High	Mobile	Centralized	Accuracy is improved	[7]
SVM	High	Static	Distributed	Accuracy is improved	[8]
	Moderate	Static	Centralized	Accuracy is improved	[9]
Bayesian	Moderate	Static	Centralized	Improved energy efficiency	[10]
	Low	Static	Centralized	Time complexity is reduced	[11]
	High	Static	Distributed	Accuracy is improved	[12]

3.2 Connectivity and Coverage

'Connectivity' indicates each and every node that sends information to a receiver through relays or directly, and includes no isolated nodes. 'Coverage' indicates monitoring and all effectively deployed area nodes. Random deployment of nodes is feasible when compared to deterministic deployment. Full coverage and partial coverage are types of coverage. Sweep, barrier, target, and focused are further classifications of partial coverage. The advantages of connectivity and coverage include coverage of the target location by an optimal number of nodes quickly and dynamically, without any lost information; additionally, dynamically changing paths either in connected or disconnected networks is also possible. Table 4 illustrates machine learning techniques for connectivity and coverage.

Table 4. Machine learning techniques for connectivity and coverage

Machine learning mechanism	Complexity	Connectivity or coverage	Network	Mobility of nodes	Contribution	Ref.
Regression	Low	Connectivity	Centralized	Static	Reliability and quality of network is optimized	[13]
SVM	Moderate	Connectivity	Distributed	Static	Efficiency is improved	[14]
Random forest	Moderate	Coverage	Distributed	Static	Accuracy is improved	[15]
Bayesian	Moderate	Coverage	Distributed	Static	Time complexity is reduced	[16]
k-means & fuzzy c-means	Low	Connectivity	Distributed	Static	Workload is reduced	[17]
Reinforcement learning	Low	Coverage	Distributed	Static	Network lifetime is improved	[18]
	Moderate	Connectivity	Centralized	Static or mobile	Network lifetime is improved	[19]

3.3 Anomaly Detection

Anomaly detection refers to an inconsistent and significant fluctuation that can appear when measuring data readings. For example, traffic monitoring in an application produces data and transmits it through relay nodes, where readings are continuously changing; data can sometimes also be lost, and as such, it must be protected by detecting data attacks. Using machine learning techniques, anomaly detection can be improved. The advantages of clustering algorithms include lowering overhead; bypass the complexity of hybrid attacks by detecting affected nodes and the type of anomaly. To manage attacks and faults in dynamic network environments, machine learning techniques are used which overcomes the problem of anomaly by considering history and modifying data parameters accordingly. Table 5 shows machine learning based anomaly detection techniques.

3.4 Fault Detection

Faults can include battery failure, communication issues, hardware failure, software failure, topological changes, and an inefficient base station. Detecting faults caused by deployment changes, resource limitations, accuracy between faulty and normal nodes, and type of surroundings represents complex obstacles. Applying machine learning approaches to detect faults delivers benefits such as the categorisation of faults and enhanced accuracy. Table 6 shows fault detection techniques using machine learning.

Table 5. Machine learning based anomaly detection techniques

Machine learning mechanism	Complexity	Anomaly	Network	Contributions	Ref.
K-NN	Low	Cyber-attacks, Random faults	Distributed	Time complexity is reduced	[20]
Decision tree	Moderate	Sinkhole problem	Centralized	Accuracy is improved	[21]
SVM	High	Detected anomaly	Centralized	Time complexity is reduced	[22]
	High	Detected intrusion	Centralized	Accuracy is improved	[23]
Bayesian	Moderate	Detected outlier	Distributed	Accuracy is improved	[24]
	High	Trust management issue	Distributed	Accuracy is improved	[25]
K-means	Moderate	Hybrid anomaly	Centralized	Accuracy is improved	[26]
Q-learning	High	Denial of service attack	Distributed	Network lifetime is improved	[27]
Regression	Moderate	Detected anomaly	Centralized	Accuracy is improved	[28]
Deep learning	High	Detected intrusion	Centralized	Accuracy is improved	[29]

Table 6. Fault detection techniques using machine learning

Machine learning mechanism	Complexity	Accuracy (%)	Diagnosing faults	Ref.
SVM	Low	99	Negative alerts	[30]
	High	98	Faulty nodes	[31]
Bayesian	Moderate	30	Body sensors	[32]
	High	100	Faulty nodes	[33]
	Moderate	98	Faulty nodes	[34]
ANN	High	98	Faulty sensors	[35]
K-NN	Moderate	99	Faulty nodes	[36]
Semi-supervised learning	Moderate	99	Detected faulty node	[37]
Deep learning	High	99	Faulty data	[38]

3.5 Routing

Transmission bandwidth, processing capacity, memory, and power affect routing in a network. The purpose of efficient routing is to find the best optimal route that consumes the least resources in terms of power, while providing an increased lifetime. The nodes near to a base station act as relay nodes that consume more power, which presents an issue for lifetime of a network. Employing machine learning techniques for wireless networks has particular benefits; for example, changes in surroundings can be adopted via machine learning without the need for reprogramming, while selecting optimum cluster heads in routing can minimise communication overheads. Additionally, by employing machine learning in applications, optimum paths can be found and latency reduced. Table 7 depicts machine learning based routing protocols.

3.6 Congestion Control

Congestion control is affected when data transmission is above the capacity of a communication channel, which can arise when sending data in several rather than one pattern, causing buffer overflows, packet collisions, contention within the transmission channel, and dynamically varying time in obtaining the packet. Congestion affects packet delivery ratio, quality of service, end-to-end latency, and energy consumption. Employing machine learning techniques for congestion control has advantages, as end-to-end delay can be minimised by finding an optimal path, and traffic can be estimated by accurately using machine learning algorithms. Table 8 illustrates congestion control strategies using machine learning techniques.

Table 7. Machine learning based routing protocols

Machine learning mechanism	Complexity	Quality of service	Network	Mobility of nodes	Topology	Ref.
ANN	High	No	Centralized	Static	Tree	[39]
	Moderate	Yes	Distributed	Static	Tree	[40]
	Moderate	No	Distributed	Mobile	Tree	[41]
Deep learning	High	No	Centralized	Mobile	Hybrid	[42]
SVM	Moderate	No	Distributed	Static	Hybrid	[43]
Bayesian	Moderate	No	Distributed	Static	Tree	[44]
	Low	No	Centralized	Static	Hybrid	[45]
	Moderate	No	Centralized and decentralized	Mobile	Hybrid	[46]
K-means	Low	Yes	Distributed	Static	Hybrid	[47]
	Moderate	No	Distributed	Static	Tree	[48]
	Moderate	No	Centralized	Static	Hybrid	[49]
Singular value decomposition	Moderate	No	Distributed	Static	Arbitrary	[50]

Table 8. Congestion control strategies using machine learning techniques

Machine learning mechanism	Parameters considered	Flow of data	Flow of control	Contributions	Ref.
ANN	Consumption of energy & end-to-end delay	Continuous	Hop by hop	Controlled traffic	[51]
Fuzzy logic	Packet loss ratio	Continuous	Hop by hop	Managing queue length	[52]
SVM	Consumption of energy, throughput & latency	Continuous	Hop by hop	Controlled rate of transmission	[53]
Reinforcement learning	Energy efficiency & throughput	Continuous	Hop by hop	Controlled traffic	[54]

3.7 Medium Access Control

Medium access control forms a data link sub-layer used in addressing channel allocation, recognising frames, and transferring data from upper layers in a network. The lifetime of a network can be enhanced through energy efficient protocols, which is critical in terms of network dynamic nature and removing noise in data. Medium access control can be schedule-based (communication among nodes, required in particular slots of time), or contention-based (central communication not required). Machine learning techniques for medium access control has advantages that include avoiding latency, reducing energy consumption, reducing extra efforts to reconstruct the network by adding nodes or removing dead nodes, minimising end-to-end delay, and enhancing the efficiency of self-learning in the network. Table 9 shows Machine learning based medium access control designs for wireless networks.

Table 9. Machine learning based medium access control (MAC) designs

Machine learning mechanism	Complexity	MAC type	Category	Synchronization	Ref.
Reinforcement learning	High	Hybrid	Contention	Yes	[55]
	Moderate	Hybrid	Schedule	Yes	[56]
	Low	ALOHA	Contention	Yes	[57]
		CSMA	Schedule	Yes	[58]
Random forest	Low	Hybrid	Contention	No	[59]

3.8 Data Aggregation

Data aggregation combines and collects data from nodes affected by memory, power, and computational and communication overheads. Efficient mechanisms of data aggregation enhance the lifetime of networks and balance energy utilisation. Tree-based, cluster-based, centralised, and in-network are types of data aggregation. Efficient cluster heads selection is effected by machine learning to balance energy consumption at nodes; furthermore, using machine learning techniques, for data aggregation adapts to dynamic environments without the need for reprogramming or reconfiguring, can reduce the dimensionality of data, thereby lowering overheads in communication at cluster heads or nodes, and in this way reduce transmission delay. Table 10 depicts the advantages of using machine learning techniques in data aggregation for wireless networks.

Table 10. Advantages of using machine learning techniques in data aggregation

Machine learning mechanism	Complexity	Network	Topology	Mobility of nodes	Improvement	Ref.
Regression	Low	Distributed	Tree	Static	Network lifetime	[60]
K-NN	Moderate	Distributed	Hybrid	Static	Network lifetime	[61]
Decision tree	Moderate	Distributed	Tree	Static	Accuracy	[62]
	Low	Distributed	Tree	Static	Network lifetime	[63]
ANN	High	Centralized	Hybrid	Static	Accuracy	[64]
Bayesian	Moderate	Distributed	Hybrid	Static or mobile	Accuracy	[65]
	Low	Centralized	Hybrid	Static	Time complexity	[66]
	Moderate	Centralized	Tree	Static	Accuracy	[67]
K-means	Moderate	Distributed	Tree	Static	Eliminated redundancy	[68]
Hierarchical clustering	Moderate	Centralized or distributed	Hierarchical	Static	Transmissions reduced	[69]
Principle component analysis	Low	Distributed	Tree	Static	Network lifetime	[70]
	High	Distributed	Tree	Static	Network lifetime	[71]
	Moderate	Distributed	Hybrid	Static	Network lifetime	[72]
Singular value decomposition	High	Centralized	Hybrid	Static	Jnnecessary transmission reduced	[73]
Genetic classifier	Moderate	Distributed	Star	Static or mobile	Network lifetime	[74]

3.9 Target Tracking

Target tracking is the monitoring and detecting of dynamic or static phenomena in a network, obtained by tracking a target via multiple nodes, thereby obtaining accurate results for single nodes is difficult. Target tracking includes a number of issues, such as missing targets, tracking latency, node failure, data aggregation, energy consumption, connectivity and coverage, which can be solved by divide and conquer techniques or estimations by prior comparative studies. Benefits include reduced overheads in tracking any mobile or stationary target, and for dynamic networks, machine learning algorithms can improve efficiency. Table 11 illustrates algorithms for node or target tracking by machine learning.

3.10 Quality of Service

Quality of service is the level of service provided by a network. This may relate to a specific application, e.g., active nodes, measurements of nodes, and deployment or network-specific aspects, e.g., bandwidth or rate of energy utilisation. It is impacted on by unbalanced traffic, dynamic networks, data redundancy, resource constraints, scalability, energy balancing, and variations in traffic type. Table 12 depicts machine learning approaches towards quality of service in wireless networks.

Table 11. Algorithms for node or target tracking by machine learning

Machine learning mechanism	Mobility of target	Number of targets	Mobility of sensor	Number of sensors	Contributions	Ref.
Bayesian & reinforcement learning	Static	Single	Static or mobile	Multiple	Network lifetime is improved	[75]
Principle component analysis	Static	Single	Static	Multiple	Network lifetime is improved	[76]
Q learning	Static	Single	Static	Multiple	Task scheduling is efficient	[77]
Genetic algorithm	Static	Single	Static	Multiple	Network lifetime is improved	[78]
Memetic algorithm	Static	Single	Static	Multiple	Network lifetime is improved	[74]
Bayesian	Static	Single	Static	Single	Accuracy is improved	[79]
	Mobile	Single	Mobile	Single	Accuracy is improved	[80]
	Static	Single	Static	Multiple	Communication overhead is reduced	[81]
	Static	Multiple	Static	Single or multiple	Network lifetime is improved	[82]

Table 12. Machine learning approaches towards quality of service

Machine learning mechanism	Complexity	Contributions	Ref.
ANN	Moderate	Detected fault nodes	[35]
	Moderate	Balancing energy	[83]
	Low	Estimated link quality	[84]
Reinforcement learning	Moderate	Communication framework for cross layers	[85]
	Low	Protocol for data dissemination & controlled topology	[86]
	High	Satisfied constraint service composition	[87]
	Moderate	Cooperative distributed adaptive routing	[88]

3.11 Synchronization

Synchronisation is used for power management, data aggregation, sleep scheduling, localisation, transmission scheduling, target tracking, and security by the protocol designs of these respective functions. One way, two-way, and receiver synchronisation are the classifications for types in this regard. Common time frames may be different for nodes within a network.

Capriglione et al. [89] consider noise, clock frequency, clock drift, and latency by regression technique for synchronisation in low-cost networks. The authors [90] propose linear regression for extended time synchronisation in dynamic, clock drift networks, which employ automatic resynchronisation, thereby reducing error and enhancing accuracy of networks. Betta et al. [91] consider clock resolution, drift, and jitter by using regression techniques in applications, thus enhancing synchronisation performance.

3.12 Event Detection

Event detection is the process of detecting misbehaving events in data when monitoring data continuously and making decisions about events. Requirements include a lower false alarm rate, limited power and computational resources, synchronisation and detection rates that are highly accurate (which are affected by memory). Using machine learning techniques in event detection has advantages such as being able to detect events from complex data, and achieving effective duty cycles by enhancing packet delivery ratio.

Illiano and Lupu [92] focus on enhancing accuracy by regression using the KNN method to effectively extract required data from raw data provided by nodes. Li et al. [93] propose query processing using the KNN method to extract the necessary information from stored information, thus reducing processing time for queries and balancing energy consumption. Han et al. [94] introduce fuzzy logic and rule-based approaches in event detection, by collecting neighbour nodes' information, thereby improving accuracy and speed.

3.13 Mobile Sink

A mobile sink gathers data from each node by visiting them directly, rather than gathering data from node to base station in multiple hops. However, in large networks, it is difficult for a mobile sink to visit each node in a rapid manner. This can be solved by introducing rendezvous points, where the mobile sink visits some nodes to gather data, while the remaining nodes transmit data to the nearest rendezvous points. To avoid delay, multiple rendezvous points can be used, but this increases network cost. Using machine learning approaches for mobile sinks has advantages, including obtaining an optimum number of rendezvous points or cluster heads, assisting in gathering data effectively, determining and selecting an optimal path by delay-aware mobile sink routing.

Wang et al. [95] introduced concepts for mobile sink and store the gathered data in cloud where every mobile sink is considered as fog device that can be a bridge between cloud and wireless networks. It collects data in parallel thus reducing delay, energy and improving scheduling, lifetime. Tashtarian et al. [96] proposed scheduling by selecting rendezvous points by optimal deadline trajectory techniques and virtual structures by decision trees. The authors [97] presented naive Bayes classifier for gathering data by mobile sink effectively than conventional approaches.

Table 13. Machine learning based techniques for energy harvesting

Machine learning mechanism	Complexity	Source of energy	Network	Ref.
Reinforcement learning	Low	Solar energy	Centralized	[77]
Deep learning	High	Solar energy	Centralized	[78]
Hierarchical clustering	Low	Solar or wind energy	Distributed	[98]
Regression	High	Solar energy	Centralized	[99]
	High	Solar energy	Centralized or distributed	[100]

3.14 Energy Harvesting

The lifetime of a network depends on the energy utilised by nodes in a network, where battery is the source of energy. Using energy harvesting techniques such as sleep scheduling, routing, a mobile charger, mobile sink, or employing efficient protocols, can prolong the lifetime of network to several years. Energy harvesting is affected by unreachable nodes, additional maintenance, and large computational resources, but long-lasting, self-powered, and maintenance free. Energy sourced from wind, radio frequencies, thermal sources, solar energy, vibration, as well as mechanical energy, are used for network energy harvesting. Harvesting can be done without storage or backup sources, and with a stored or powered, rechargeable battery. Using machine learning techniques for energy harvesting has several advantages, e.g., the amount of energy harvested within a given time can be estimated, thereby improving network performance. Table 13 depicts machine learning based techniques for energy harvesting.

Table 14. Machine learning algorithms to solve various challenges in wireless networks

Issues in wireless networks	Machine learning mechanism	Contributions
Localization	Reinforcement learning	It works for dynamic changes in network without any initial knowledge
	K-NN	Estimated distance efficiently in free range network localization
Connectivity & coverage	Decision tree	Classification for isolated and connected network nodes performed efficiently
	Deep learning & evolutionary computation	Optimal connectivity & Coverage can be done with minimum nodes
Fault and anomaly detection	Random forest	Classifying normal and fault nodes
	Principle component analysis	Anomaly is detected
	Independent component analysis	Anomaly is detected
	Deep learning	Detected faults & Online anomaly
Routing	Decision tree, random forest & evolutionary computation	Routing optimally is predicted for controlling data by alternate dynamic paths
Medium access control	SVM	Assigned channels efficiently
	Decision tree	Assigned channels efficiently
	Deep learning	Predicted slots in time & reconfigure to dynamic network automatically
Data aggregation	K-means	Optimum number of cluster heads are found
	SVM	Optimum number of cluster heads are found
	Reinforcement learning	Routing paths are selected optimally without any initial knowledge
Congestion control	Reinforcement learning	Congestion is determined & alternate paths are found
	Random forest, decision tree & SVM	In large scale wireless networks, normal and congestion nodes classification is performed
	Evolutionary computation	Avoided congestion by selecting alternative dynamic optimal path
	Principle & independent component analysis	Unnecessary information sending is controlled by reducing dimensionality
Target tracking	Deep learning	Tracking targets in multiple mobile networks efficiently
	SVM	Targets classification in heterogeneous networks
	Decision tree	Targets classification in heterogeneous networks
Event Detection	Principle & independent component analysis	Event detection from complex data
	Evolutionary computation	Duty cycles are managed effectively
	Deep learning	Duty cycles are managed effectively
Mobile Sink	Evolutionary computation	Selecting best path between rendezvous points or sensor nodes from mobile sink
	Reinforcement learning	Selecting optimal tour & rendezvous points
	Random forest	Forwarding routes for data & selecting rendezvous points optimally in large scale networks
Energy harvesting	SVM	Energy harvested in a particular time is estimated
	Deep learning	Energy harvested in a particular time is estimated
	Evolutionary computation	Energy harvested amount is estimated
Synchronization	Deep learning	Allocating channels & dynamical resynchronization

4. Statistical Analysis and Limitations

4.1 Statistical Analysis

In this section recent research overview on machine learning algorithms are presented. Table 14 illustrates machine learning algorithms to solve various challenges in wireless networks.

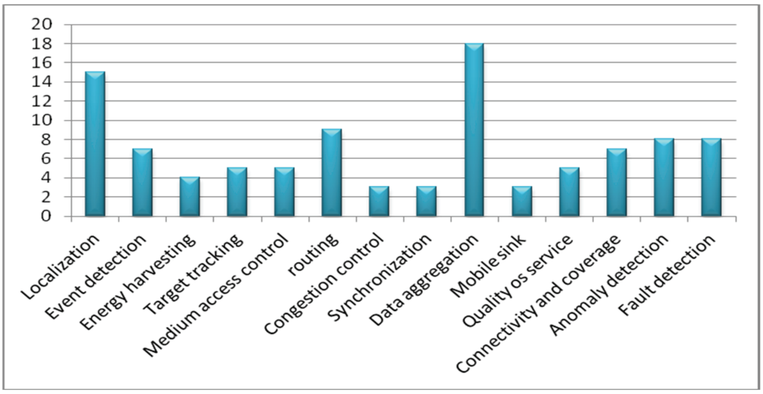

Fig. 2. Statistical charts for issues.

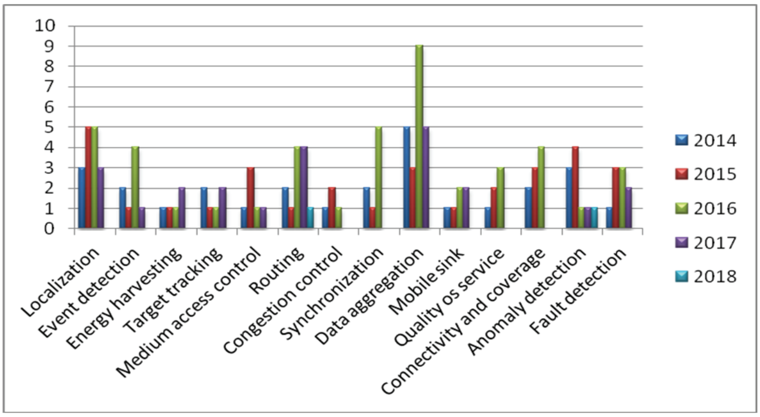

Fig. 3. Research papers published year wise.

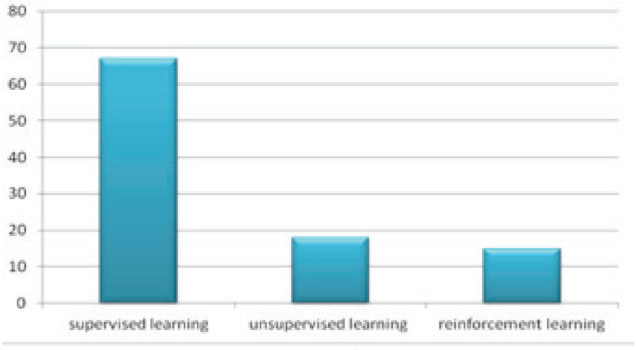

Fig. 4. Classification based on learning technique.

Figs. 2 and 3 shows wireless networks issues where many researchers are interested. Fig. 4 estimates percentage of issues solved by machine learning techniques by its classification. Fig. 5 estimates percentage of issues solved by machine learning algorithms.

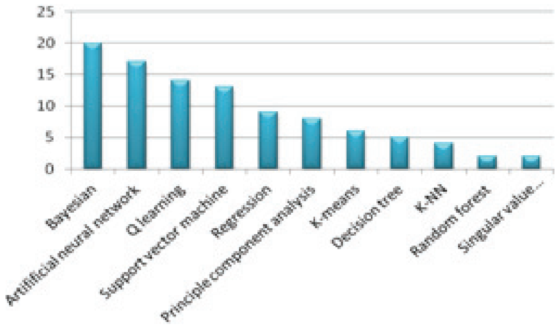

Fig. 5. Issues solved by machine learning based algorithms.

4.2 Limitations

Despite its advantages, applying machine learning algorithms has some limitations. (1) Predicting accurate results immediately is not possible, as learning from history is required for machine learning algorithms. (2) Historical data amounts determine performance, but the consumption of energy is high if large amounts of data require processing. There is thus a trade-off between computational overheads and energy consumption, but this can be solved by running machine learning algorithms centrally. (3) It is difficult to draw predictions from algorithms that require validation in real world applications. (4) Identifying an efficient machine learning technique to address a particular challenge can at times be difficult.

5. Open Challenges

In this section we will provide open challenges in wireless networks which need further research to solve those challenges by employing effective machine learning techniques. Defining standards and designing cross layer protocols for quality of service in heterogeneous networks with different requirements is difficult and challenging task where further research to be done to standardize the quality of service. Table 15 shows the key issues need to be addressed in wireless networks.

- **Localisation:** It is initially important to find an optimal path via beacon nodes within a network, but to our knowledge, there is no predefined strategy for planning a path in a network. We explored machine learning in two-dimensional space for localisation accuracy, with less energy utilisation; however, localisation in three-dimensional space for dynamic and static networks need to be conducted, which is necessary in most applications in real time.

- **Connectivity and coverage:** To cover the target area using a lower number of nodes, and placing them effectively, is a challenging task. Random deployment of nodes in real-time applications leads to the coverage hole problem in dynamic environments. In this situation, finding accurate solutions in three-dimensional spaces, and providing suitable algorithms with lower computational overheads, requires further research.

- **Anomaly detection:** Many researchers have proposed various techniques for detecting anomalies that affect transmission delay, communication overheads, and misleading data. However, future research is required to overcome the anomaly once it has been detected in order to lower damages of network. Speed and accuracy in detecting and selecting anomaly techniques remains an open challenge that needs to be improved for heterogeneous networks.

- **Routing:** Transmitting data packets from sender to receiver in one-to-one communication is easy, but many-to-many is difficult due to packet collisions, which remains an ongoing issue for further research aimed at overcoming collisions. Implementing effective protocols for dynamic network changes that occur due to external factors is required.

- **Data aggregation:** This is a simple procedure for uniform data rates, but for non-uniform data rates, further research is needed to reduce complexity. Using a mobile sink can enhance energy efficiency; however, scheduling a mobile sink is a challenging issue for non-uniform data rates. Scalability, energy efficiency, and lower costs are factors that need to be considered for data aggregation.

- **Congestion control and avoidance:** Data loss, as well as internal and external factors, lead to congestion in dynamic networks, where implementing congestion control techniques are needed to lower data transmission rates. Efficient, fast, effective strategies are needed, alongside self-learning within dynamic networks to remove or add nodes accordingly when congestion arises. Estimating traffic rates in fast dynamic networks via efficient protocols remains an ongoing research issue, as does collecting data, in addition to sending it among nodes to be determined.

- **Energy harvesting:** To enhance the lifetime of wireless energy harvesting systems with limited power, lower costs and high efficiency is required for synchronisation between cross layers, i.e., medium access control and physical layers. Machine learning strategies for improving reliability in large-scale networks that adapt to dynamic changes via self-charging and discharging duty cycles is an additional research direction of interest.

- **Quality of service:** Quality of service is required to meet the needs of applications and users in determining data rates, and for handling traffic, costs of network, energy consumption.

Table 15. Key issues need to be addressed in wireless networks

Challenges in wireless networks	Key issues need to be addressed
Localization	Localization in real time applications with three dimensional spaces for dynamic and static networks needs to be performed.
Connectivity and coverage	Deployment of nodes in three dimensional spaces with less computational overheads in real time applications needs to be performed.
Anomaly detection	Speed and accuracy in detecting and selecting anomaly techniques is a key issue which needs to be improved for heterogeneous networks.
Routing	Implementing effective protocols and reducing packet collisions for dynamic network changes caused by external factors is a key issue which needs to be determined.
Data aggregation	Data aggregation is complex for non-uniform data rates where further research to be done to reduce complexities.
Congestion control and avoidance	To estimate traffic rates in fast dynamic network by efficient protocols as well as collecting data apart from sending data among nodes to be determined to control and avoid congestion.
Energy harvesting	Machine learning strategies to improve reliability in large scale networks that adapts to dynamic changes by self-charging and discharging duty cycles is a key issue which needs to be performed.
Quality of service	Defining standards and designing cross layer protocols in heterogeneous networks with different requirements is difficult, further research to be done to standardize the quality of service.

6. Conclusion

Several machine learning techniques for wireless networks were discussed in this survey. We addressed various factors such as fault detection, localisation, anomaly detection, data aggregation, routing, synchronisation, medium access control protocols, selecting a mobile sink path, energy harvesting, quality of service, and congestion control, all of which can be addressed by employing machine learning techniques. Additionally, we compared and produced statistical reports of various machine learning techniques' impact on wireless networks. In this survey, we suggest selecting a particular machine-learning technique to address a challenge in wireless networks. Finally, ongoing issues for further research in future were presented.

References

[1] M. Bernas and B. Placzek, "Fully connected neural networks ensemble with signal strength clustering for indoor localization in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 11, no. 12, article no. 403242, 2015.

[2] X. Li, S. Ding, and Y. Li, "Outlier suppression via non-convex robust PCA for efficient localization in wireless sensor networks," *IEEE Sensors Journal*, vol. 17, no. 21, pp. 7053-7063, 2017.

[3] W. Zhao, S. Su, and F. Shao, "Improved DV-hop algorithm using locally weighted linear regression in anisotropic wireless sensor networks," *Wireless Personal Communications*, vol. 98, no. 4, pp. 3335-3353, 2018.

[4] S. S. Banihashemian, F. Adibnia, and M. A. Sarram, "A new range-free and storage-efficient localization algorithm using neural networks in wireless sensor networks," *Wireless Personal Communications*, vol. 98, no. 1, pp. 1547-1568, 2018.

[5] A. El Assaf, S. Zaidi, S. Affes, and N. Kandil, "Robust ANNs-based WSN localization in the presence of anisotropic signal attenuation," *IEEE Wireless Communications Letters*, vol. 5, no. 5, pp. 504-507, 2016.

[6] S. Phoemphon, C. So-In, and D. T. Niyato, "A hybrid model using fuzzy logic and an extreme learning machine with vector particle swarm optimization for wireless sensor network localization," *Applied Soft Computing*, vol. 65, pp. 101-120, 2018.

[7] N. Baccar and R. Bouallegue, "Interval type 2 fuzzy localization for wireless sensor networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, article no. 42, 2016.

[8] T. Tang, H. Liu, H. Song, and B. Peng, "Support vector machine based range-free localization algorithm in wireless sensor network," in *Machine Learning and Intelligent Communications*. Cham: Springer, 2016, pp. 150-158.

[9] J. Hong and T. Ohtsuki, "Signal eigenvector-based device-free passive localization using array sensor," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 4, pp. 1354-1363, 2015.

[10] T. L. T. Nguyen, F. Septier, H. Rajaona, G. W. Peters, I. Nevat, and Y. Delignon, "A Bayesian perspective on multiple source localization in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1684-1699, 2015.

[11] B. Sun, Y. Guo, N. Li, and D. Fang, "Multiple target counting and localization using variational Bayesian EM algorithm in wireless sensor networks," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 2985-2998, 2017.

[12] Y. Guo, B. Sun, N. Li, and D. Fang, "Variational Bayesian inference-based counting and localization for off-grid targets with faulty prior information in wireless sensor networks," *IEEE Transactions on Communications*, vol. 66, no. 3, pp. 1273-1283, 2017.

[13] W. Sun, X. Yuan, J. Wang, Q. Li, L. Chen, and D. Mu, "End-to-end data delivery reliability model for estimating and optimizing the link quality of industrial WSNs," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 3, pp. 1127-1137, 2017.

[14] W. Kim, M. C. Stankovic, K. H. Johansson, and H. J. Kim, "A distributed support vector machine learning over wireless sensor networks," *IEEE Transactions on Cybernetics*, vol. 45, no. 11, pp. 2599-2611, 2015.

[15] W. Elghazel, K. Medjaher, N. Zerhouni, J. Bahi, A. Farhat, C. Guyeux, and M. Hakem, "Random forests for industrial device functioning diagnostics using wireless sensor networks," in *Proceedings of 2015 IEEE Aerospace Conference*, Big Sky, MT, 2015, pp. 1-9.

[16] B. Yang, Y. Lei, and B. Yan, "Distributed multi-human location algorithm using naive Bayes classifier for a binary pyroelectric infrared sensor tracking system," *IEEE Sensors Journal*, vol. 16, no. 1, pp. 216-223, 2016.

[17] J. Qin, W. Fu, H. Gao, and W. X. Zheng, "Distributed k-means algorithm and fuzzy c-means algorithm for sensor networks based on multiagent consensus theory," *IEEE Transactions on Cybernetics*, vol. 47, no. 3, pp. 772-783, 2017.

[18] H. Chen, X. Li, and F. Zhao, "A reinforcement learning-based sleep scheduling algorithm for desired area coverage in solar-powered wireless sensor networks," *IEEE Sensors Journal*, vol. 16, no. 8, pp. 2763-2774, 2016.

[19] E. Ancillotti, C. Vallati, R. Bruno, and E. Mingozzi, "A reinforcement learning-based link quality estimation strategy for RPL and its impact on topology management," *Computer Communications*, vol. 112, pp. 1-13, 2017.

[20] M. Xie, J. Hu, S. Han, and H. H. Chen, "Scalable hypergrid k-NN-based online anomaly detection in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 8, pp. 1661-1670, 2013.

[21] A. Garofalo, C. Di Sarno, and V. Formicola, "Enhancing intrusion detection in wireless sensor networks through decision trees," in *Dependable Computing*. Heidelberg: Springer, 2013, pp. 1-15.

[22] Z. Feng, J. Fu, D. Du, F. Li, and S. Sun, "A new approach of anomaly detection in wireless sensor networks using support vector data description," *International Journal of Distributed Sensor Networks*, vol. 13, no. 1, article no. 1550147716686161, 2017.

[23] H. S. Emadi and S. M. Mazinani, "A novel anomaly detection algorithm using DBSCAN and SVM in wireless sensor networks," *Wireless Personal Communications*, vol. 98, no. 2, pp. 2025-2035, 2018.

[24] C. Titouna, M. Aliouat, and M. Gueroui, "Outlier detection approach using Bayes classifiers in wireless sensor networks," *Wireless Personal Communications*, vol. 85, no. 3, pp. 1009-1023, 2015.

[25] R. Feng, X. Han, Q. Liu, and N. Yu, "A credible Bayesian-based trust management scheme for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 11, no. 11, article no. 678926, 2015.

[26] M. Wazid and A. K, Das, "An efficient hybrid anomaly detection scheme using K-means clustering for wireless sensor networks," *Wireless Personal Communications*, vol. 90, no. 4, pp. 1971-2000, 2016.

[27] S. Shamshirband, A. Patel, N. B. Anuar, M. L. M. Kiah, and A. Abraham, "Cooperative game theoretic approach using fuzzy Q-learning for detecting and preventing intrusions in wireless sensor networks," *Engineering Applications of Artificial Intelligence*, vol. 32, pp. 228-241, 2014.

[28] S. Haque, M. Rahman, and S. Aziz, "Sensor anomaly detection in wireless sensor networks for healthcare," *Sensors*, vol. 15, no. 4, pp. 8764-878, 2015.

[29] T. Ma, F. Wang, J. Cheng, Y. Yu, and X. Chen, "A hybrid spectral clustering and deep neural network ensemble algorithm for intrusion detection in sensor networks," *Sensors*, vol. 16, no. 10, article no. 1701, 2016.

[30] S. Zidi, T. Moulahi, and B. Alaya, "Fault detection in wireless sensor networks through SVM classifier," *IEEE Sensors Journal*, vol. 18, no. 1, pp. 340-347, 2017.

[31] M. Jiang, J. Luo, D. Jiang, J. Xiong, H. Song, and J. Shen, "A cuckoo search-support vector machine model for predicting dynamic measurement errors of sensors," *IEEE Access*, vol. 4, pp. 5030-5037, 2016.
[32] H. Zhang, J. Liu, and N. Kato, "Threshold tuning-based wearable sensor fault detection for reliable medical monitoring using Bayesian network model," *IEEE Systems Journal*, vol. 12, no. 2, pp. 1886-1896, 2018.
[33] C. Titouna, M. Aliouat, and M. Gueroui, "FDS: fault detection scheme for wireless sensor networks," *Wireless Personal Communications*, vol. 86, no. 2, pp. 549-562, 2016.
[34] W. Meng, W. Li, Y. Xiang, and K. K. R. Choo, "A Bayesian inference-based detection mechanism to defend medical smartphone networks against insider attacks," *Journal of Network and Computer Applications*, vol. 78, pp. 162-169, 2017.
[35] P. Chanak and I. Banerjee, "Fuzzy rule-based faulty node classification and management scheme for large scale wireless sensor networks," *Expert Systems with Applications*, vol. 45, pp. 307-321, 2016.
[36] W. Li, P. Yi, Y. Wu, L. Pan, and J. Li, "A new intrusion detection system based on KNN classification algorithm in wireless sensor network," *Journal of Electrical and Computer Engineering*, vol. 2014, article no. 240217, 2014.
[37] M. Zhao and T. W. Chow, "Wireless sensor network fault detection via semi-supervised local kernel density estimation," in *Proceedings of 2015 IEEE International Conference on Industrial Technology (ICIT)*, Seville, Spain, 2015, pp. 1495-1500.
[38] G. Gennarelli and F. Soldovieri, "Performance analysis of incoherent RF tomography using wireless sensor networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2722-2732, 2016
[39] A. Mehmood, Z. Lv, J. Lloret, and M. M. Umar, "ELDC: an artificial neural network based energy-efficient and robust routing scheme for pollution monitoring in WSNs," *IEEE Transactions on Emerging Topics in Computing*, 2017. http://doi.org/10.1109/TETC.2017.2671847.
[40] M. S. Gharajeh and S. Khanmohammadi, "DFRTP: dynamic 3D fuzzy routing based on traffic probability in wireless sensor networks," *IET Wireless Sensor Systems*, vol. 6, no. 6, pp. 211-219, 2016.
[41] J. R. Srivastava ad T. S. B. Sudarshan, "A genetic fuzzy system based optimized zone based energy efficient routing protocol for mobile sensor networks (OZEEP)," *Applied Soft Computing*, vol. 37, pp. 863-886, 2015.
[42] Y. Lee, "Classification of node degree based on deep learning and routing method applied for virtual route assignment," *Ad Hoc Networks*, vol. 58, pp. 70-85, 2017.
[43] F. Khan, S. Memon, and S. H. Jokhio, "Support vector machine based energy aware routing in wireless sensor networks," in *Proceedings of 2016 2nd International Conference on Robotics and Artificial Intelligence (ICRAI)*, Rawalpindi, Pakistan, 2016, pp. 1-4.
[44] V. Jafarizadeh, A. Keshavarzi, and T. Derikvand, "Efficient cluster head selection using Naïve Bayes classifier for wireless sensor networks," *Wireless Networks*, vol. 23, no. 3, pp. 779-785, 2017.
[45] Z. Liu, M. Zhang, and J. Cui, "An adaptive data collection algorithm based on a Bayesian compressed sensing framework," *Sensors*, vol. 14, no. 5, pp. 8330-8349, 2014.
[46] F. Kazemeyni, O. Owe, E. B. Johnsen, and I. Balasingham, "Formal modeling and analysis of learning-based routing in mobile wireless sensor networks," in *Integration of Reusable Systems*. Heidelberg: Springer, 2014, pp. 127-150.
[47] R. El Mezouary, A. Choukri, A. Kobbane, and M, El Koutbi, "An energy-aware clustering approach based on the k-means method for wireless sensor networks," in *Advances in Ubiquitous Networking*. Singapore: Springer, 2015, pp. 325-337.
[48] A. Ray and D. De, "Energy efficient clustering protocol based on K-means (EECPK-means)-midpoint algorithm for enhanced network lifetime in wireless sensor network," *IET Wireless Sensor Systems*, vol. 6, no. 6, pp. 181-191, 2016.
[49] B. Jain, G. Brar, and 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*. Singapore: Springer, 2018, pp. 113-123

[50] P. Guo, J. Cao, and X. Liu, "Lossless in-network processing in WSNs for domain-specific monitoring applications," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2130-2139, 2017.
[51] M. A. Alsheikh, S. Lin, D. Niyato, and H. P. Tan, "Rate-distortion balanced data compression for wireless sensor networks," *IEEE Sensors Journal*, vol. 16, no. 12, pp. 5072-5083, 2016.
[52] A. A. Rezaee and F. Pasandideh, "A fuzzy congestion control protocol based on active queue management in wireless sensor networks with medical applications," *Wireless Personal Communications*, vol. 98, no. 1, pp. 815-842, 2018.
[53] M. Gholipour, A. T. Haghighat, and M. R. Meybodi, "Hop-by-hop congestion avoidance in wireless sensor networks based on genetic support vector machine," *Neurocomputing*, vol. 223, pp. 63-76, 2017.
[54] S. H. Moon, S. Park, and S. J. Han, "Energy efficient data collection in sink-centric wireless sensor networks:a cluster-ring approach," *Computer Communications*, vol. 101, pp. 12-25, 2017.
[55] M. Rovcanin, E. De Poorter, I. Moerman, and P. Demeester, "A reinforcement learning based solution for cognitive network cooperation between co-located, heterogeneous wireless sensor networks," *Ad Hoc Networks*, vol. 17, pp. 98-11, 2014.
[56] I. Mustapha, B. M. Ali, A. Sali, M. F. A. Rasid, and H. Mohamad, "An energy efficient reinforcement learning based cooperative channel sensing for cognitive radio sensor networks," *Pervasive and Mobile Computing*, vol. 35, pp. 165-184, 2017.
[57] S. Kosunalp, Y. Chu, P. D. Mitchell, D. Grace, and T. Clarke, "Use of Q-learning approaches for practical medium access control in wireless sensor networks," *Engineering Applications of Artificial Intelligence*, vol. 55, pp. 146-154, 2016.
[58] K. H. Phung, B. Lemmens, M. Goossens, A. Nowe, L. Tran, and K. Steenhaut, "Schedule-based multi-channel communication in wireless sensor networks: a complete design and performance evaluation," *Ad Hoc Networks*, vol. 26, pp. 88-102, 2015.
[59] B. Alotaibi and K. Elleithy, "A new mac address spoofing detection technique based on random forests," *Sensors*, vol. 16, no. 3, article no. 281, 2016.
[60] X. Song, C. Wang, J. Gao, and X. Hu, "DLRDG: distributed linear regression-based hierarchical data gathering framework in wireless sensor network," *Neural Computing and Applications*, vol. 23, no. 7-8, pp. 1999-2013, 2013.
[61] Y. Li and L. E. Parker, "Nearest neighbor imputation using spatial–temporal correlations in wireless sensor networks," *Information Fusion*, vol. 15, pp. 64-79, 2014.
[62] F. Edwards-Murphy, M. Magno, P. M. Whelan, J. O'Halloran, and E. M. Popovici, "b+ WSN: smart beehive with preliminary decision tree analysis for agriculture and honey bee health monitoring," *Computers and Electronics in Agriculture*, vol. 124, pp. 211-219, 2016.
[63] H. He, Z. Zhu, and E. Makinen, "Task-oriented distributed data fusion in autonomous wireless sensor networks," *Soft Computing*, vol. 19, no. 8, pp. 2305-2319, 2015.
[64] C. Habib, A. Makhoul, R. Darazi, and C. Salim, "Self-adaptive data collection and fusion for health monitoring based on body sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2342-2352, 2016.
[65] A. De Paola, P. Ferraro, S. Gaglio, G. L. Re, and S. K. Das, "An adaptive Bayesian system for context-aware data fusion in smart environments," *IEEE Transactions on Mobile Computing*, vol. 16, no. 6, pp. 1502-1515, 2017.
[66] S. Hwang, R. Ran, J. Yang, and D. K. Kim, "Multivariated Bayesian compressive sensing in wireless sensor networks," *IEEE Sensors Journal*, vol. 16, no. 7, pp. 2196-2206, 2015.
[67] Y. Wang, A. Yang, Z. Li, X. Chen, P. Wang, and H. Yang, "Blind drift calibration of sensor networks using sparse Bayesian learning," *IEEE Sensors Journal*, vol. 16, no. 16, pp. 6249-6260, 2016.
[68] H. Harb, A. Makhoul, and R. Couturier, "An enhanced K-means and ANOVA-based clustering approach for similarity aggregation in underwater wireless sensor networks," *IEEE Sensors Journal*, vol. 15, no. 10, pp. 5483-5493, 2015.

[69] X. Xu, R. Ansari, A. Khokhar, and A. V. Vasilakos, "Hierarchical data aggregation using compressive sensing (HDACS) in WSNs," *ACM Transactions on Sensor Networks (TOSN)*, vol. 11, no. 3, article no. 45, 2015.
[70] A. Morell, A. Correa, M. Barcelo, and J. L. Vicario, "Data aggregation and principal component analysis in WSNs," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3908-3919, 2016.
[71] M. I. Chidean, E. Morgado, E. del Arco, J. Ramiro-Bargueno, and A. J. Caamano, "Scalable data-coupled clustering for large scale WSN," *IEEE Transactions on Wireless Communications*, vol. 14, no. 9, pp. 4681-4694, 2015.
[72] M. Wu, L. Tan, and N. Xiong, "Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications," *Information Sciences*, vol. 329, pp. 800-818, 2016.
[73] A. R. Pinto, C. Montez, G. Araujo, F. Vasques, and P. Portugal, "An approach to implement data fusion techniques in wireless sensor networks using genetic machine learning algorithms," *Information Fusion*, vol. 15, pp. 90-101, 2014.
[74] S. N. Das, S. Misra, B. E. Wolfinger, and M. S. Obaidat, "Temporal-correlation-aware dynamic self-management of wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2127-2138, 2016.
[75] P. Braca, P. Willett, K. LePage, S. Marano, and V. Matta, "Bayesian tracking in underwater wireless sensor networks with port-starboard ambiguity," *IEEE Transactions on Signal Processing*, vol. 62, no. 7, pp. 1864-1878, 2014.
[76] B. Zhou, Q. Chen, T. Li, and P. Xiao, "Online variational Bayesian filtering-based mobile target tracking in wireless sensor networks," *Sensors*, vol. 14, no. 11, pp. 21281-21315, 2014.
[77] B. Xue, L. Zhang, W. Zhu, and Y. Yu, "A new sensor selection scheme for Bayesian learning based sparse signal recovery in WSNs," *Journal of the Franklin Institute*, vol. 355, no. 4, pp. 1798-1818, 2018.
[78] H. Chen, R. Wang, L. Cui, and L. Zhang, "Easidslt: a two-layer data association method for multitarget tracking in wireless sensor networks," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 1, pp. 434-443, 2015.
[79] 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," *Sensors and Actuators B: Chemical*, vol. 224, pp. 266-274, 2016.
[80] Z. Wei, Y. Zhang, X. Xu, L. Shi, and L. Feng, "A task scheduling algorithm based on Q-learning and shared value function for WSNs," *Computer Networks*, vol. 126, pp. 141-149, 2017.
[81] M. Elhoseny, A. Tharwat, A. Farouk, and A. E. Hassanien, "K-coverage model based on genetic algorithm to extend WSN lifetime," *IEEE Sensors Letters*, vol. 1, no. 4, pp. 1-4, 2017.
[82] C. P. Chen, S. C. Mukhopadhyay, C. L. Chuang, T. S. Lin, M. S. Liao, Y. C. Wang, and J. A. Jiang, "A hybrid memetic framework for coverage optimization in wireless sensor networks," *IEEE Transactions on Cybernetics*, vol. 45, no. 10, pp. 2309-2322, 2015.
[83] M. Collotta, G. Pau, and A. V. Bobovich, "A fuzzy data fusion solution to enhance the QoS and the energy consumption in wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2017, article ID 341828, 2017.
[84] W. Sun, W. Lu, Q. Li, L. Chen, M. Mu, and X. Yuan, "WNN-LQE: wavelet-neural-network-based link quality estimation for smart grid WSNs," *IEEE Access*, vol. 5, pp. 12788-12797, 2017.
[85] E. K. Lee, H. Viswanathan, and D. Pompili, "RescueNet: reinforcement-learning-based communication framework for emergency networking," *Computer Networks*, vol. 98, pp. 14-28, 2016.
[86] A. P. Renold and S. Chandrakala, "MRL-SCSO: multi-agent reinforcement learning-based self-configuration and self-optimization protocol for unattended wireless sensor networks," *Wireless Personal Communications*, vol. 96, no. 4, pp. 5061-5079, 2017.
[87] L. Ren, W. Wang, and H. Xu, "A reinforcement learning method for constraint-satisfied services composition," *IEEE Transactions on Services Computing*, 2017. http://doi.org/10.1109/TSC.2017.2727050.

[88] M. A. Razzaque, M. H. U. Ahmed, C. S. Hong, and S. Lee, "QoS-aware distributed adaptive cooperative routing in wireless sensor networks," *Ad Hoc Networks*, vol. 19, pp. 28-42, 2014.
[89] D. Capriglione, D. Casinelli, and L. Ferrigno, "Analysis of quantities influencing the performance of time synchronization based on linear regression in low cost WSNs," *Measurement*, vol. 77, pp. 105-116, 2016.
[90] J. J. Perez-Solano and S. Felici-Castell, "Adaptive time window linear regression algorithm for accurate time synchronization in wireless sensor networks," *Ad Hoc Networks*, vol. 24, pp. 92-108, 2015.
[91] G. Betta, D. Casinelli, and L. Ferrigno, "Some notes on the performance of regression-based time synchronization algorithms in low cost WSNs," in *Sensors*. Cham: Springer, 2015, pp. 439-443.
[92] V. P. Illiano and E. C. Lupu, "Detecting malicious data injections in event detection wireless sensor networks," *IEEE Transactions on Network and service management*, vol. 12, no. 3, pp. 496-510, 2015.
[93] Y. Li, H. Chen, M. Lv, and Y. Li, "Event-based k-nearest neighbors query processing over distributed sensory data using fuzzy sets," *Soft Computing*, vol. 23, no. 2, pp. 483-495, 2019.
[94] Y. Han, J. Tang, Z. Zhou, M. Xiao, L. Sun, and Q. Wang, "Novel itinerary-based KNN query algorithm leveraging grid division routing in wireless sensor networks of skewness distribution," *Personal and Ubiquitous Computing*, vol. 18, no. 8, pp. 1989-2001, 2014.
[95] T. Wang, J. Zeng, Y. Lai, Y. Cai, H. Tian, Y. Chen, and B. Wang, "Data collection from WSNs to the cloud based on mobile Fog elements," *Future Generation Computer Systems*, 2017. https://doi.org/10.1016/j.future.2017.07.031.
[96] F. Tashtarian, M. Y. Moghaddam, K. Sohraby, and S. Effati, "ODT: optimal deadline-based trajectory for mobile sinks in WSN: a decision tree and dynamic programming approach," *Computer Networks*, vol. 77, pp. 128-143, 2015.
[97] S. Kim and D. Y. Kim, "Efficient data-forwarding method in delay-tolerant P2P networking for IoT services," *Peer-to-Peer Networking and Applications*, vol. 11, no. 6, pp. 1176-1185, 2018.
[98] S. W. Awan and S. Saleem, "Hierarchical clustering algorithms for heterogeneous energy harvesting wireless sensor networks," in *Proceedings of 2016 International Symposium on Wireless Communication Systems (ISWCS)*, Poznan, Poland, 2016, pp. 270-274.
[99] A. Sharma and A. Kakkar, "Forecasting daily global solar irradiance generation using machine learning," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 2254-2269, 2018.
[100] W. M. Tan, P. Sullivan, H. Watson, J. Slota-Newson, and S. A. Jarvis, "An indoor test methodology for solar-powered wireless sensor networks," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 16, no. 3, article no. 82, 2017.

Nimmagadda Srilakshmi https://orcid.org/0000-0002-4414-0174

She is pursuing as a PhD Research Scholar in School of Computer Science and Engineering (SCOPE) in Vellore Institute of Technology (VIT), Vellore, India. She received M.Tech. degree and achieved gold medal from Gitam University, Hyderabad, India. Her current research work includes machine learning, cloud computing, wireless ad hoc and sensor networks and mobile communications.

Arun Kumar Sangaiah https://orcid.org/0000-0002-0229-2460

He is working as an Associate Professor in School of Computer Science and Engineering (SCOPE) in Vellore Institute of Technology (VIT), Vellore, India. His current research work includes global software development, wireless ad hoc and sensor networks, machine learning, cognitive networks and advances in mobile computing and communications.