



Article Recent Studies Utilizing Artificial Intelligence Techniques for Solving Data Collection, Aggregation and Dissemination Challenges in Wireless Sensor Networks: A Review

Walid Osamy ^{1,†}, Ahmed M. Khedr ^{2,3,†}, Ahmed Salim ^{3,4,†} and Amal Ibrahim AlAli ^{5,†} and Ahmed A. El-Sawy ^{1,*,†}

- ¹ Computer Science Department, Faculty of Computers and Artificial Intelligence, Benha University, Benha 13513, Egypt; walid.osamy@fci.bu.edu.eg
- ² Computer Science Department, University of Sharjah, Sharjah 27272, United Arab Emirates; akhedr@sharjah.ac.ae
- ³ Mathematics Department, Zagazig University, Zagazig 44523, Egypt; ahmed.salimxpress@gmail.com or a.salem@qu.edu.sa
- ⁴ Department of Computer Science, College of Sciences and Arts, Qassim University, P.O. Box 931, Buridah 51931, Saudi Arabia
- ⁵ Information Systems Department, University of Sharjah, Sharjah 27272, United Arab Emirates; aialali@sharjah.ac.ae
- Correspondence: ahmed.el_sawy@fci.bu.edu.eg
- † These authors contributed equally to this work.

Abstract: The growing importance and widespread adoption of Wireless Sensor Network (WSN) technologies have helped the enhancement of smart environments in numerous sectors such as manufacturing, smart cities, transportation and Internet of Things by providing pervasive realtime applications. In this survey, we analyze the existing research trends with respect to Artificial Intelligence (AI) methods in WSN and the possible use of these methods for WSN enhancement. The main goal of data collection, aggregation and dissemination algorithms is to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced. In this paper, we highlight data collection, aggregation and dissemination challenges in WSN and present a comprehensive discussion on the recent studies that utilized various AI methods to meet specific objectives of WSN, during the span of 2010 to 2021. We compare and contrast different algorithms on the basis of optimization criteria, simulation/real deployment, centralized/distributed kind, mobility and performance parameters. We conclude with possible future research directions. This would guide the reader towards an understanding of up-to-date applications of AI methods with respect to data collection, aggregation and dissemination challenges in WSN. Then, we provide a general evaluation and comparison of different AI methods used in WSNs, which will be a guide for the research community in identifying the mostly adapted methods and the benefits of using various AI methods for solving the challenges related to WSNs. Finally, we conclude the paper stating the open research issues and new possibilities for future studies.

Keywords: artificial intelligence; data aggregation; Internet of Things; data collection; data dissemination; challenges; wireless sensor networks

1. Introduction

The field of ad hoc network technology has experienced remarkable research attention over recent years [1]. Ad hoc networks are broadly classified into two categories: Mobile ad hoc Network (MANET) and Wireless Sensor Network (WSN). WSNs have limited power consumption and are comprised of low cost devices when compared to MANETs [2,3]. MANETs are designed for mobile devices that are capable of moving freely in any direction independently, while WSN nodes with embedded CPU and smart sensors are generally deployed and used for data sensing and monitoring of surrounding environmental factors



Citation: Osamy, W.; Khedr, A.M.; Salim, A.; AlAli, A.I.; El-Sawy, A.A. Recent Studies Utilizing Artificial Intelligence Techniques for Solving Data Collection, Aggregation and Dissemination Challenges in Wireless Sensor Networks: A Review. *Electronics* 2022, *11*, 313. https:// doi.org/10.3390/electronics11030313

Academic Editors: Federico Alimenti and Hirokazu Kobayashi

Received: 24 November 2021 Accepted: 5 January 2022 Published: 20 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). such as wind, air, humidity, pressure, vibration, detecting gases and chemicals, earthquake and so on [2,4]. The increasing benefits of IoT applications in daily life have revolutionized people's lifestyle choices. For effective communication of data among nodes, many such IoT-based applications demand exact identification of node position and location [5]. WSN is regarded as the core component of IoT and enables a diverse set of applications [6–10]. The integration of WSN into the IoT allows the dynamic connection of sensor devices to the internet and achieves the tasks assigned. WSN provides valuable capabilities to applications such as in science, military, engineering, area monitoring, forest fire detection, disaster prediction and so on, but at the same time faces different challenges due to its constrained resources and capabilities [11].

The key purpose of this survey is to give insight into up-to-date applications of Artificial Intelligence (AI) methods with respect to different WSN challenges. We highlight various challenges related to Data Collection (DC), aggregation and dissemination in WSNs and present a comprehensive discussion of the recent studies that utilize various AI methods to meet specific objectives of WSN, during the span of 2010 to 2021. Then, we provide a general evaluation and comparison of different AI methods used in WSNs, which will be a guide for the research community in identifying the most adapted methods to address the different WSN challenges, and the benefits of using various AI methods for solving each of these challenges related to WSNs.

In this survey, we review most of the existing AI techniques and their applications in WSNs for overcoming the DC, Aggregation and Dissemination challenges of WSNs. We primarily present an overview of these WSN challenges and a review of various AI techniques. Then, we present the applications of AI techniques for solving these challenges and enhancing WSN performance in terms of packet transfer rate, WSN lifetime, energy efficiency etc. This study will provide the reader with adequate understanding about the challenging concerns of WSN and the power of AI techniques in solving the WSN challenges.

While survey papers exist on analyzing different challenges facing WSNs, most of them focused on applying AI methods to solve a particular problem, such as routing or energy usage, while others focused on solving some of the challenges faced by WSN. There has been related work that discussed or partially surveyed the literature related to AI based protocols and algorithms solving different WSN challenges. The work in this survey is different from others as our intention is to provide a contemporary survey of the more recent literature. Our focus is to compare different AI techniques which allow to explore new strategies for resolving existing WSN problems and to enhance the WSN performance, and we cover several optimization techniques that address different challenges in WSNs related to DC, Aggregation and Dissemination. The core content of this survey is AI schemes for WSN while we cover all aspects of it in addressing these WSN challenges. Moreover, we have reviewed the papers published from 2010 to 2021 to carry out a systematic analysis and comparison. The paper provides a comprehensive survey of 55 relevant papers across different academic disciplines such as AI and computer science from reliable database sources.

The methodology used to conduct this survey is discussed in the following subsection.

1.1. Research Methodology

The research methodology used here is split into four parts, Articles selection, Articles classifications, Article analysis and Discussion and future scope. The articles selection phase (Phase 1) is comprised of two steps: database source selection and article selection & filtering. The classification of articles is covered in Phase 2. Phase 3 involves article analysis (discussed later in Section 4) along with some open research challenges. Finally, Phase 4 includes the discussion and future scope which is covered later in the Section 5 where we conclude the paper.

• Database sources selection step: Paper sources and key search parameters might have an impact on the quality of research. To achieve this, the papers in this review are taken from reliable databases: Web of Science (https://clarivate.com/webofsciencegroup/ solutions/web-of-science/, accessed on 1 June 2021), Scopus (https://www.scopus. com/, accessed on 1 June 2021), IEEE Explorer (https://ieeexplore.ieee.org/, accessed on 1 June 2021), and ACM digital library (https://dl.acm.org/, accessed on 1 June 2021). Moreover, only indexed journals are considered. The query search string and search keywords connected to the study subject are carefully chosen in order to conduct a good search that covers the most relevant material.

 Articles selection and filtering step: The search queries are made up of researchrelated words, brief phrases, and Boolean operators (ANDs and ORs).
 We carried out search by selecting one of the AI keywords with WSN keywords, along with WSN or Wireless sensor network as a keyword. The search results are limited to journal articles and the predefined range of years. Then all the search results are combined and filtered.

The diagram in Figure 1 shows the whole processes of generating the query strings. Query search strings in abstracts, keywords, and titles are run on the specified databases to locate the core relevant publications from 2010 to 2021. Moreover, journal articles indexed in more than one database are included and other types are excluded. The resultant search terms are combined and filtered to choose the primary publications, excluding those that are not directly linked to our subject, duplicated, or of insufficient quality. Moreover, in order to determine the eligibility of the filtered articles to our targets first abstract is read and if it doesn't contain an indicator of its eligibility then the content of the article is investigated. Otherwise, the article is selected. Using this, 55 relevant research articles are chosen as primary papers where 13% of articles are selected from IEEE/ACM journals and 87% of articles are selected from articles. Figure 2 shows the number of selected journal research articles per year, along with the number of related and review articles per year (note that the selection criteria are not applied on the selection of related review paper).



Figure 1. Query search string formation diagram.



Figure 2. Number of selected research articles and review articles per year from 2010 to 2021.

1.1.2. Articles Classifications Phase

Among 55 papers evaluated, various AI methods in WSNs have been recognized and classified from primary database sources as shown in Figure 3. These techniques include Fuzzy Logic, Artificial Neural Network, Evolutionary Computation, Nature-inspired, Multi-Agent Systems, Trajectory based, Physical computation, Reinforcement Learning and Hybrid. An overview of these techniques is presented later in Section 2.2. Moreover, this classification of AI techniques is employed during the discussion of DC, Aggregation and Dissemination challenges in WSN to show how AI techniques handled each challenge.



Figure 3. AI methods applied to address Data Collection, Aggregation and Dissemination WSN challenges.

1.2. Paper Contributions

Among the papers reviewed, various AI techniques in WSNs have been checked and classified. An overview of these techniques is presented initially. This classification of AI techniques is then employed during the discussion of each challenge in WSN to show how AI techniques handled each challenge. Some papers may cover multiple aspects and will be surveyed for each category. The paper analyses the research distribution and trends that

characterize the use of AI in WSN. In addition, the paper identifies challenges, promising research directions in applying AI-based solutions to various WSN challenges, with the aim to promote and facilitate further research. In summary, the following are the paper's major contributions:

- 1. We review the existing AI techniques and their applications in WSNs to overcome the challenge issues of WSNs.
- 2. We present an overview of the major challenges in WSNs and the various AI techniques to handle DC, Aggregation and Dissemination challenges.
- 3. A comprehensive discussion on the recent studies that utilized various AI methods to meet specific objectives of WSN during the span of 2010 to 2021 is given.
- 4. We present a solid comparison between the used AI techniques for solving each of these challenges.
- 5. We identify promising research directions in applying AI-based solutions to various WSN challenges, with the aim to promote and facilitate further research.

1.3. Paper Organization

The rest of the article is structured as follows: Background Information discussed in Section 2. Section 3 examines related research works. In Section 4, Data collection, aggregation and dissemination challenges are discussed and corresponding AI solutions are presented, along with the respective open research issues that can guide the research immunity for future innovations. Section 5 provides a discussion and concludes the paper. Table 1 summarizes the abbreviations for the key terms used in the paper.

Abbreviation	Description
ANN	Artificial Neural Network.
ACO	Ant Colony Optimization.
ABC	Artificial Bee Colony.
AI	Artificial Intelligence.
BA	Bat Algorithm.
CI	Computational Intelligence.
DC	Data Collection.
DL	Deep Learning.
DNN	Deep Neural Network.
GA	Genetic Algorithm.
IoT	Internet of Things.
PSO	Particle Swarm Optimization.
RL	Reinforcement Learning.
SI	Swarm Intelligence.
WSN	Wireless Sensor Network.

Table 1. Table of main Abbreviations.

2. Background Information

In this section, a brief background information on the focus of our survey. Here, we will first discuss DC, Aggregation and Dissemination challenges in WSNs, followed by different AI techniques for WSNs.

2.1. Data Collection, Aggregation and Dissemination

In WSNs, the data gathering and dissemination process aims to gather data from nodes and convey it to the sink/ BS in such a way that the highest data rate is accomplished

while the total WSN lifespan is maximized. The data might be plain raw data or data that has been processed using basic techniques like aggregation or compression. The act of aggregating data from various nodes and sending it to the BS for further analysis is known as data aggregation. WSN gathers data from environment for a long duration and direct the delivery of collected data from nodes to sink (or BS), which may contain highly redundant data and increases the communication load. Data coming from different sensors are redundant and huge for the BS to process. Data aggregation is carried out to deal with this problem and the BS obtains only the significant information. The key purpose of data fusion is to reduce duplication of data from multiple sources and also decrease the transmissions' count and conserve resources [12]. On the other hand, data dissemination refers to the process for distributing the data and queries for data to be routed in network. Source node in WSN is considered as node that generates data in relation to an event, and node that is interested in data is considered as the sink node [13]. Event is information that needs to be reported while interest refers to the descriptors for events that node is interested in.

2.2. Artificial Intelligence Techniques

AI refers to a computer system's ability to accomplish activities that need human intelligence while imitating human minds or ideas. It is considered as one of the important area of computer science that tries to make machine "smart". The most widely used techniques of AI include: search methods, learning methods, fuzzy system, knowledge based and reasoning. AI is applicable to resolve many complex concerns in various sectors such as security, finance, health care, transportation and so on, through its ability in handling deficient and noisy data, dealing with non-linear issues, and are suitable for use in prediction and faster generalization once trained [14]. We have studied and classified various AI techniques used in WSN as depicted in Figure 3. Different AI techniques used in addressing WSN challenges include Fuzzy Logic, Artificial Neural Network, Evolutionary Computation, Nature-inspired, Multi-Agent Systems, Trajectory based, Physical Computation, Swarm Intelligence, Deep Learning, Reinforcement Learning and Hybrid. An overview of these techniques is presented as follows.

1. Metaheuristics

Metaheuristics are the most common type of algorithms that use a degree of randomness to achieve optimal solutions to hard problems (or as optimal as possible) [15]. Metaheuristics are applied to a large number of areas. Metaheuristic algorithms can be categorized in various ways. For example, one scheme of classification is: trajectorybased and population-based approaches [16]. Trajectory-based schemes typically aim to locate a single optimal solution through piecewise style movement in the design (search) space (e.g., simulated annealing). While population-based schemes use multiple solution through search space and cooperate with each other to reach the final solution (e.g., evolutionary computation, physical inspired computation and nature inspired computation). Evolutionary computation is inspired by biological evolution and natural selection, crossover or recombination and mutation (e.g., Genetic Algorithm, Differential Evolution and Memetic Algorithm). Physical inspired computation is inspired by physical areas such as classical and quantum mechanics, thermodynamics, electromagnetism, relativity, and optics [17] (e.g., Central Force Optimization, Gravitational Search Algorithm, Intelligent Water Drops and so on). Nature inspired computations imitate colonies, birds, flocks, insects in their living method or individuals communication (e.g., harmony, bat algorithm, cuckoo search etc.). The collective behavior that arises from a group of social insects has been called Swarm Intelligence (SI). SI deals with the cooperation of numerous homogeneous individuals in the environment [18]. Such techniques involve strategies and share information among the individuals for self-organization, learning and co-evolution during iterations to provide high efficiency. The individuals follow very simple rules and as there is no central infrastructure available to show how individuals behave, interaction can take place between individuals and these individuals as a population can exchange related data using any message-carrier [19]. Multiple interacting intelligent agents can solve a problem that is hard to solve by an individual agent or monolithic system, by searching and interacting with environment. Agents search for other neighboring agents and interact with them or with the environment to learn new things and to make decisions. To complete their assigned mission, agents utilize their knowledge, make decisions and conduct actions in the environment [20].

2. Learning Methods

One of the most important feature in the human (or animals) is learning. Learning is the ability to automatically acquire new information and improve it via experience without requiring any explicit programming. So, learning is a part of AI like Artificial Neural Network (ANN), Reinforcement Learning (RL) and Deep Learning (DL).

With the ability to mimic biological neural network and human attributes, ANNs have been successful in solving complex challenging problems. ANN consists of small interconnected devices known as nodes inspired from the biological neurons in a brain. Information is passed from these interconnected devices using links represented by an arrow. Input and weight are the two values associated with an incoming connection, whose summation will generate the unit's output. After training an ANN using training data sets, new data sets can be introduced so that the trained ANN can be used further for prediction and classification purposes. The key advantage of using ANNs over other methods lies in its ability to model non-linear and complex processes without much interruption between input and output variables. It is used to solve many problems related to prediction and validation, optimization, function approximation, clustering, time series analysis and pattern recognition. Several architectures of ANN are present in literature which include: Radial Basis Function network, Multi-Layer Perception (MLP), Back-Propagation and Recurrent Neural Network (RNN) [14].

RL is a branch of AI concerned with how intelligent agents should interact in a given environment in order to maximize the concept of cumulative reward. Learning is accomplished by interaction between learning entities and their surrounding environment in the RL process. Objects attempt to learn through trial and error. The value function, the environment, and the reinforcement function are the three main components of RL. The RL environment is generally dynamic, with a range of possible states. There is a set of viable actions for each condition at any given time [21].

With the ability to learn without human supervision, drawing from data that is both unstructured and unlabeled, DL is an attractive AI function based on representation learning. DL mimics the human brain in processing data and creates patterns that are used in decision making. The architecture of DL includes several layers in between input and output layers and non-linear information processing units. DL is considered as a universal learning scheme and it is used to find solution to all kinds of problems in various application areas [22]. DL is also used to solve problems of big data analytics which include determining the volume of input information necessary to represent DL algorithms and obtaining good data abstractions and representations [23]. Feature extraction is represented in multiple hierarchical levels, which distinguishes DL from other machine learning approaches. DL is used in various situations where machine intelligence is useful:

- People can't explain their expertise (sound and speech recognition, language understanding and vision).
- The solution needs adaptation to a particular case (e.g., personalization, biometrics).
- If the solution to a problem changes over time (e.g., stock and price prediction, tracking, weather prediction).
- Human expert is absent (e.g., navigation on Mars).
- The problem size is huge for limited reasoning capabilities (e.g., finding matching ADs to Facebook, calculating webpage ranks, sentiment analysis).

Currently, DL is practically used in almost every field. Hence, this technique is often termed as Universal Learning Technique [22].

3. Fuzzy Logic (FL)

FL is another AI technique that imitates the way of human decision making. It is used for uncertain reasoning or managing incomplete information [14]. The possibilities are either True (T) or False (F). FL works on the basis of 'truth-value' between 0 and 1 [24]. Fuzzy set membership can take any value between 0 and 1. Examples include centroid defuzzification, maximum and mean-of-maxima [21].

These are some of the recent AI techniques that have been popularly used in addressing the DC, Aggregation and Dissemination challenges. In the upcoming sections, we present an up-to-date survey of existing research trends with respect to AI methods applied in WSN. Then, we go through the major challenges of WSN that includes: DC, aggregation and dissemination; and identify the different methods to address these challenges through the use of different AI techniques along with open research issues in each of these categories. We further present a comprehensive discussion on the recent studies regarding these techniques. Moreover, we compare them to determine suitable technique(s) in overcoming various challenging concerns of WSN and state the open research issues and future research directions.

3. Related Works

Over the years, the research in the field of WSN is becoming more active and numerous works have been done for WSN enhancement. To address the challenges of WSN, different surveys are conducted based on various factors. Some of the recent research in solving challenges of WSN using AI techniques are presented here. Kulkarni et al. [25] conducted a survey of solving various WSN challenges using Computational Intelligence (CI) techniques. They showed that many CI algorithms have outperformed under severe and uncertain environment conditions and limited power supply, however, some solutions are not the best as no real time evaluations of these solutions are done on practical WSN scenarios. According to their findings, design and deployment challenges are usually centralized issues and for this, Artificial Neural Networks (ANN), Genetic Algorithm and Particle Swarm Optimization are very much suited. While SI is considered as a good paradigm for routing in MANETS, the large communication overhead makes it necessary that the SI models have to be modified to suit the properties and requirements of WSN.

Another survey related to SI techniques in WSN is presented by Sarobin et al. [26]. They provided some recent analysis and showed that energy efficiency can be enhanced by considering the following challenges of WSN which include: node localization, scheduling, design and deployment, clustering and data aggregation. The survey briefly explains how SI methods are utilized to solve these challenges and improve WSN lifetime, coverage and connectivity. They also illustrate that most of the works are on stationary WSNs, so mobile WSNs must be encouraged in future. The majority of existing SI-based approaches rely solely on simulation. They should be implemented and tested in real-time environment.

A discussion on Intelligent Optimization of WSN using bio-inspired computing can be found in the survey conducted by Jabbar et al. [27]. They presented the solutions for non-biological systems using bio-inspired algorithms. Moreover, they presented that some of the issues of WSN can be solved using hybrid approaches of CI. Routing issues can be solved by SI as it possess scalable, robust, adaptive, and distributed properties. SI includes variety of algorithms such as Fish school, Bee colony, Ant Colony Optimization (ACO) algorithms, GA , PSO, and ANN. Run-time of ANN and GA are longer when compared to others. FL is fast and deterministic algorithm but optimization is an issue. Artificial Immune System is better when compared to GA but its computational run time is very long. They provided a foundation for the future researches to aim to look for hybrid approaches that optimize the memory and computational power of WSN.

Montoya et al. [28] focused on main issues of WSN such as coverage, data fidelity, connectivity, and WSN lifetime. In this work, they presented the principles and algorithms of AI to optimize the issues of WSN. They presented a MAS based approach which comprises of a group of agents interacting with one another. Agents are responsible for perceiving the environment through sensor devices and respond to the environment using actuators. The associated field of AI that covers the principles of MAS is termed as Distributed AI. Solarium SunSPOT emulator is used to test the SunSPOT devices without the requirement of any additional hardware platform. However, this system is not tested on real WSN scenario, so this needs to be done in future.

Yu et al. [29] presented a survey on Intelligent techniques in WSN for minimizing energy consumption. AI techniques are applied over WSN to aggregate the data and to reduce redundancy through optimization of routing protocols for preserving the energy requirements. GA does not require complex computations, so it is suitable for application having limited computing power. While there exist various research works in implementing energy efficient algorithms, they tried to implement sensors which are intelligent enough to present optimal solutions.

Maksimovi et al. [30] presented a survey related to the use of FL in WSN. FL is considered as a promising approach to evaluate diverse parameters in an efficient manner. It improves decision making, reduces resource consumption and is suitable for resolving WSN issues such as routing, data aggregation, security, localization as well as deployment. FL can tolerate unreliable or imprecise readings easily. They presented it to be an easier and efficient technique. The disadvantage of using this technique is that the rules count grows exponentially with the variables count and storage requires much extra memory. They showed that the FL approach can solve the shortcomings of most of the algorithms. As this is a rule based approach and due to constant traversal of rules, it may slow down the event detection and decision making process. To solve this problem they also presented rule based reduction techniques which are efficient, but none of them can be considered as a general solution. A key property that must be kept in mind is that the reduction technique is very challenging. Future work that needs to be done is to implement a software based FL to enhance the speed of the system.

The work in [31] focuses on the most recent advancements in data aggregation techniques for WSN concerning terrestrial, underground, underwater areas by providing a review of different data aggregation strategies and protocols. However, the role of AI approaches in resolving data aggregation challenges in WSNs has been described in a very limited way, without deep discussion. Ref. [12] gives a methodical literature review of data aggregation in the field of WSNs. The study covers data aggregation strategies, tools, methodology, and challenges. However, the focus of this study was not on the role of AI methods in data aggregation, but rather on data aggregation methodologies using different criteria. Ref. [32] focuses on data aggregation approaches based on routing protocols. The work in [33] reviews the important contributions to data aggregation security solutions that primarily employ soft computing methodologies. Protocols are categorized according to soft computing approaches such as FL, SI, GA, and ANNs. Discussions on metaheuristics and how to employ them to solve the deployment problem are addressed in [34], and how to use them to solve lifetime problems of WSN are covered in [35]. Multi-Objective Optimization techniques in the context of WSNs are presented in [36] to introduce the development efforts for surveillance and monitoring. Other sophisticated optimization strategies were explored together with the needs of different optimization approaches such as mathematical programming and heuristics/meta-heuristics based optimization. A survey of ACO techniques for static and mobile WSNs is proposed in [37].

An overview of a set of aggregation techniques ranging from Space Filling Curves, to Q-digest, Wavelets, Gossip aggregation, and Compressive Sensing presented in [38]. In [39] a study reviews various data aggregation techniques such as clustered aggregation, tree-based aggregation, in-network aggregation, and centralized data aggregation with focus on energy consumption of sensor nodes. In [40] data aggregation approaches systematically reviewed, examined and categorized into Flat Network and Hierarchical Network

based on different factors (e.g., types of networks and algorithms, node heterogeneity, and the aggregator mobility). In-network data aggregation algorithms are surveyed and analyzed in [41] based on various parameters like energy efficiency, aggregation rate, network lifetime, number of alive nodes, throughput, energy-saving techniques, scalability, network security, route overlap, and route repair mechanisms. A review of DC techniques proposed in [42] with objectives to providing a detailed comparison between the maximum amount shortest path (MASP) and zone-based energy-aware (ZEAL) DC protocols and to help researchers in selecting the most suitable protocol for a WSN application. Ref. [43] surveyed traditional WSN then provide some discussions on Power Consumption and Data Aggregation for WSN.

The core content of this survey is the use of AI schemes for WSN while we cover all aspects of it in addressing DC, Aggregation and Dissemination challenges in WSN. In order to carry out systematic review and comparison, we have reviewed the papers published from 2010 to 2021.

We have outlined a number of open research questions that need to be addressed in the future. We hope that this article, with its rich bibliography content, will provide valuable insight into recent trends in research on the WSN and encourage new research.

4. Data Collection, Aggregation and Dissemination Challenges in WSNs

Data aggregation refers to the process of combining or gathering data coming from various sources and aggregating them to remove redundancies and improves transmission, thus saving energy of a system. In traditional approaches, packets are transferred through pre-defined routes and are aggregated on some specific nodes. Such kind of topologies consume more energy in constructing and maintaining the WSN. So, we should avoid these challenges. Table 2 summarizes the AI-based solutions to Data Aggregation and Dissemination problems in WSN.

4.1. AI Based Solutions to Data Collection, Aggregation and Dissemination Challenges in WSNs

Several AI techniques are applied to handle the DC, aggregation and dissemination challenges in WSN. We discuss them as follows:

A family of ACO based optimization algorithms introduced in [44] for data aggregation is called Data Aggregation ACO Algorithm (DAACA). It consists of three phases. First is initialization, followed by packet transmission and then, the operations on pheromones are done. They also designed evaporation and deposition of pheromones which take the advantage of both local and global merits. Four AC heuristic algorithms are implemented in this work, Basic-DAACA, Maximum and Minimum based DAACA (MM-DAACA), Elitist Strategy based DAACA (ES-DAACA), Ant Colony System based DAACA (ACS-DAACA). Basic-DAACA is the elementary algorithm on the basis of which other three algorithms are designed. The other three algorithms are designed for improving the Basic-DAACA. ES-DAACA attaches extra global pheromone concentration into Basic-DAACA. MM-DAACA restricts the pheromone range in ES-DAACA. ACS-DAACA includes both local as well as global pheromone modifications to Basic-DAACA. Basic-DAACA eliminates the repetition of creating and preserving the WSN topology which results in preventing computational overhead and prolongs network lifetime by saving energy. DAACA family has no additional energy expenditure for creating and maintaining of the topology of the network. The pheromone adjustment will not cost much energy. The average energy cost of these four methods follow the order: ACS < MM < ES < Basic. The enhanced versions of DAACA improve the quality of output of the Basic DAACA and produce data aggregation routes which are more energy conservative. The success ratio of one hop communication for DAACA is higher when compared to other algorithms. Time complexity of DAACA is lower than others for maintaining the network topology. DAACA is scalable, robust and fault tolerant. Time complexity for constructing the topology of DAACA is equivalent to ACA, and the complexity is $O(n^2)$.

The researchers have applied different swarm intelligence algorithms for solving optimization problems. One such designed technique that imitates the dynamics of river systems is the intelligent water drops algorithm. Each drop symbolizes a solution in this population-based method, and the sharing of drops during the search results in better drops (or solutions). The Intelligent Water Drops (IWDs) technique is used in [45] to generate the best data aggregation trees for WSNs. They improved the fundamental IWD method by altering the selection probability for ideal aggregation nodes, which improved the tree creation. An ant inspired solution is proposed in [46] for Minimumpower Multiresolution Data Dissemination problem. They proposed a meta-heuristic solution known as ACO based Minimum Incremental Dissemination Tree (AMIDT), with an aim to optimize the energy consumption of WSN. This algorithm includes: (1) an online tree formation algorithm and (2) Two lightweight tree adaption heuristics named path-quality and reference-rate based heuristic. Depending on the local information of the path, path-quality based heuristic was estimated, while reference based heuristic is based on aggregation rates and decreases the required count of tree reconfigurations. AMIDT exhibits improved performance in terms of energy and cost for searching a path when compared to other approaches.

ABC algorithm is another class of algorithms that showed superior performance on many kinds of optimization problems. In [47], the concept of Sparse WSN is introduced, which a special type of network characterized by the node's geographical sparsity in which the nodes cannot directly interact with each other or with other nodes, i.e., sink. One of the important concern in this paper is to reduce energy utilization on mobile robot data gathering. Therefore, the travel path length of mobile robot has to be determined to ensure high energy efficiency. This approach is similar to traveling salesman problem (TSP) with neighborhoods. So to address this, they designed an ABC-based route planning technique. Proposed ABC algorithm involves four main phases: initialization, the employed Bee, the onlooker Bee, and the scout Bee phase. The performance evaluation is done by comparing average path length with other algorithms. ABC-based approach is highly stable than other greedy approaches and converges rapidly. In case of multiple mobile robots, this approach will not work well, so this must be considered in future by implementing new scheme with joint consideration of mobile robots and travel path determination.

As the data collected by the WSN grew exponentially with increase in the number of sensors, the centralized data mining approach at the fusion center face difficulties in reducing the load and saving the WSN transmission power usage. To handle this challenge, the work in [48] proposed a distributed method of data mining for WSNs using Deep Neural Network (DNN). Unlabeled data collected from dispersed WSN nodes can be used to train the scheme's internal representations. DNN improves data mining efficiency owing to characteristics like self-learning, internal data representation and analysis, and multilayer perceptron creation with more than one hidden layer. The DNN is split into various layers distributed among the WSN nodes. This lowers the computation cost and energy consumption when compared to centralized processing. The results show an improved performance in data mining and is suitable for application in large scale WSNs.

Another memetic meta-heuristic designed for finding solution to combinatorial optimization problems is the Shuffled Frog-Leaping Algorithm (SFLA). The work in [49] adopted SFLA for the selection of optimal clusters in WSN. Cluster Head (CH) is selected based on residual energy level of the WSN nodes. Shuffled frog algorithm shows good performance, high efficiency and stronger search capability. This technique uses the shuffling method of frogs for information exchange among local searches, to achieve global optimal output. Simulations are performed by changing the nodes count to measure the count of clusters formed, energy usage, average packet loss, end to end delay, and lifetime computation.

Ref. [50] proposes a collaborative data aggregation technique that can address realtime needs in large-scale and complex networks. The PSO method is used to find the best transmission path that minimizes the maximum hop count while reducing the maximum next hop distance between two nodes. This contributes to the reduction of overall WSN energy usage. To calculate the best routing tree and meet the least energy consumption constraint, the developed PSO approach employs a multi-objective adaptive function.

A method named Hybridized Pareto-Glowworm Swarm Optimization and Authenticated Data Communication (HPGSO-ADC) is presented in [51] for reliable and secure communication of data. The CH election is based on mobility and nodes' energy. After the clustering process, the data is aggregated from various nodes to minimize the bandwidth usage and forwarded to BS using ABC approach. Authentication is done between CH and BS to guarantee the security. Various performance parameters are examined. It includes packet delivery ratio, bandwidth, energy consumption, end to end delay, throughput, WSN lifetime, availability, reliability, loss probability, and serviceability.

Every node in the WSN is expected to gather local measurements that are likely to be influenced by noise, formulate a local decision regarding the presence or absence of an event, and then send that decision to the fusion center. The fusion center then makes an ultimate decision based on specific local decisions as well as the fusion rule, therefore a successful fusion rule decision is critical. A linear decision fusion model for WSNs is suggested in [52]. To set the parameters of the linear decision fusion model, the constrained PSO method is used, and the typical penalty function is used to tackle the constrained PSO issue. Sharing the WSN infrastructure to accomplish concurrent requests is becoming a trend today, where a comparatively complicated request can be met by aggregating complementary features provided by contiguous nodes in a certain area of the network. As a solution, a mechanism namely Multi-Requests Services Optimization (MRSO) is presented in [53]. MRSO is a combinational optimization approach for aggregating multi-requests. The PSO method is used in MRSO. MRSO discovers the best solutions by sharing global best solutions across several requests. This approach improves the shareability of WSN resources across concurrent requests while also lowering network energy consumption, especially when there is a lot of temporal, geographical, and functional overlap between concurrent requests. A cluster based data aggregation and routing technique is presented in [54] for improving lifetime and throughput of a network. A heuristic algorithm called Ant Lion Optimization (ALO) is introduced to deal with the clustering problem. Improper clustering often leads to individual nodes and when such an isolated individual node transfer information to BS, it requires high transmission power that will affect the network lifetime. ALO is applied over the selected CH to yield an optimal path. This reduces the overall traveling distance of MS and also extends the network lifetime.

A Hybrid approach combining ACO and PSO for clustering and tree based data aggregation is introduced in [55] named as ACO and PSO based energy efficient clustering (ACOPSO). In this, the initial clusters are created based on residual energy, then ACOPSO is applied to enhance inter-cluster data aggregation. In ACOPSO, ACO based path selection is implemented as the primary step in which shortest path is decided with reference to the least cost based spanning tree that is created between CH and BS/ sink node. Then PSO is applied to reduce the path cost further. The performance is evaluated according to stability period, residual energy, network lifetime, and throughput. Throughput of this method is significantly better than others. WSN lifetime is improved and the energy dissipation is also consistent and balanced. The comparison is done with ABC and GA. Residual energy and throughput is considered for comparison and the results indicate a significant improvement in throughput and energy efficiency.

A mobile data acquisition technique based on clustering which uses PSO and Space Division Multiple Access (SDMA) is proposed in [56]. The proposed approach is developed to cover energy hole and offer high efficiency, minimized latency, lower energy utilization and buffer overflow reduction. SDMA is utilized for data gathering. PSO algorithm is then used to select anchor points for scheduling visiting locations for Mobile data collector. Performance parameters that are used for evaluation include: packet delivery ratio, delay incurred, throughput, and average consumed energy. Mobile Data Transporter (MDT) is a special MS node that is introduced in [57]. MDT visits and gathers information from each WSN node and transfers the required data to BS. A discrete firefly algorithm based optimization is introduced to shorten the travel distance of MDT. The scheme is then analyzed and compared with other tree based schemes and it shows that this strategy is better than others in terms of minimizing tour lengths under different conditions.

In [58], Two-Tier Distributed FL Based Protocol (TTDFP) is used to prolong multi-hop WSNs lifetime by making use of the efficiency of both clustering and routing stages together. TTDFP is an adaptive distributed protocol that operates efficiently for WSN applications and is scalable. Moreover, they utilized an optimization model to adjust the parameters of fuzzy clustering phase so as to optimize the WSN performance.

The MDF-FBCH technique of [59] performs data fusion based on the network path parameters instead of performing fusion on the data gathered by the nodes. It includes a Multisensor Data Fusion (MDF) scheme which executes fusion on the gathered parameters of the network in collaboration with Fuzzy based model for CH Selection (FBCHS). The MDF method utilized ANFIS fuzzy model to decide about the selection of network path. Instantaneous Channel State Information (I-CSI), bandwidth, Packet Loss Ratio (PLR), latency, centrality and Signal to Noise Ratio (SNR) are the crucial parameters applied to the ANFIS-based MDF model. Each parameter has its own significance in selecting the route from CH or nodes to the BS. The used parameters are then trained using multi-objective PSO algorithm. The fusion on network parameters is performed by MDF-FBCHS in mainly four steps using: a normalizer, fuzzifier, MDF engine and a defuzzifier as shown in Figure 4. The authors highlight that network factor fusion for route selection also plays a significant role in selecting an acceptable data forwarding path.



Figure 4. FBCHS fuzzy model for CH selection [59].

Chicken Swarm Optimization (CSO) Algorithm is applied to optimize the Compressive Sensing matrix in [60] that is to enhance data aggregation in cluster-based WSNs. In [61], a Simulated Annealing (SA) based algorithm for tree construction and scheduling (SATC) is proposed which aggregates data from nodes with a collision-free schedule. SATC considered average time delay in delivering the aggregated results to sink as SA's fitness function and hence provided enhanced efficiency in minimizing average latency in data aggregation than the existing scheduling techniques.

DC without delay is a significant process in WSN and is important for a variety of applications where precise actions depend on the appropriate timeline, such as event-based mission-critical applications. The authors of [62] proposed a GA based algorithm (ETDMA-GA) for efficient scheduling to minimize latency in collecting data. ETDMA-GA minimizes the communication latency by using a 2D encoding scheme for slot allocation and a fitness function to minimize the overall network latency.

A hesitant fuzzy entropy based technique for data fusion in WSN proposed in [63]. The work aims to minimize the gathering of redundant data from the WSN nodes, and exploits the information obtained from redundant data to enhance data reliability. Hesitant

fuzzy entropy method is utilized to combine the data collected from cluster nodes at the sink for gaining better data quality.

In [64], a cross layer mechanism is presented for congestion management which is enabled by FL and Oppositional ABC Optimization scheme (FCOABC). The proposed cross layer technique adopted FL with fuzzy descriptors in terms of neighboring nodes count, communication link reliability, and nodes residual energy, for selecting the CH. The WSN is arranged to form dissimilar sized clusters to address hot spots issue. In this method, smaller clusters are taken into consideration, as CHs of such clusters are at proximate locations from master station. The advantage is that they experience only reduced intra-cluster congestion and hence the relay traffic can be efficiently transferred using their preserved amount of energy. The benefit of using Oppositional ABC is that it can perform inter-cluster routing efficiently in multi-hop manner from CH to BS. Therefore, this method helps to achieve energy aware and reliable data delivery. The performance is analyzed on the basis of energy consumption by CHs, amount of data transferred to master station and total consumed energy.

The work in [65] presented a technique called GTAC-DG for DC using MS in WSN. To achieve a better trade-off between power consumption and transmission delay in the WSN, they proposed an effective rendezvous point selection process using the distributed game theoretical approach (GTBRP) and specify the suitable efficient route for sink node using the modified ACO based algorithm. The method favors nodes with high residual energy and load, which will help to efficiently balance the load and conserve energy. To avoid node buffer overflow difficulties, they included traffic load as an extra parameter to ACO and devised an efficient MS trajectory. This enables DC without traveling great distances, resulting in a better balance between DC delay and energy consumption, extending the WSN lifetime. To create a two-hop load-balanced data aggregation routing tree in the network, Ref. [66] proposed an effective method called Cuckoo-search Clustering with Two-hop Routing Tree (CC-TRT). CC-TRT employs an improved energy-aware cuckoosearch algorithm to identify the best CH for each cluster. The cuckoo-search algorithm allows for round-by-round rotation of the CH job across various sensors. The CC-TRT model is then modified to provide two new techniques: Cuckoo-search Clustering with Multi-Hop Routing (CC-MRT) and Cuckoo-search Clustering with Weighted Multi-hop Routing Tree (CC-WMRT). As an advantage, the suggested schemes offer better energy balancing among nodes in WSN.

The problem of optimizing data dissemination and routing is examined in [67]. In the first instance, the suggested scheme provides a clustering mechanism focused on the functional characteristics of the nodes based on their past behavior. Watchful Node takes the responsibility of coordinating the nodes and monitors the activities. This strategy organizes nodes into different clusters and chooses CH based on a fuzzy and location-based approach. Moreover, for each cluster, the collective behavior of the wise and sharp-minded optimizer inspired by Harris hawks pick at least one watchful Node. Then, data dissemination and routing mechanism work according to a query/event-based hybrid scheme. This maintains the stability and reliability of the WSN. The lightweight diffusion mechanism is assured by the iterative and repetitive event-based strategy, which isolates the event zone to regulate the amount of copies of data packets.

Ref. [68] suggested Group Search Optimization technique with ANN for the development of a new query-based framework for data aggregation. The scheme includes a Querying Order model that uses the Query Order ranked in terms of throughput and latency. The objective of the paper is to develop a Querying Order model to increase throughput and reduce delay, thereby improving existing schemes for data aggregation. It will allow the administrator of the network to acquire information about the suitable queries to increase the sink performance.

A data fusion scheme for Heterogeneous WSNs based on the PSO and Extreme Learning Machine (PSO-ELM) has been suggested in [69]. It optimizes the input weight matrix and hidden layer bias of the extreme learning machine using PSO. The performance

weight matrix is intended to minimize the number of hidden layer nodes and improve the generalization capabilities of the model. The sensor nodes' original data is merged via the data fusion model. The extreme learning machine was used to analyze the data obtained by the nodes in the hierarchical routing using the spatial temporal similarity between the data acquired by the heterogeneous WSN nodes.

A fuzzy-oriented geographic routing protocol called FGAF-CDG, based on the hexagonal virtual grid architecture of the GAF protocol, is proposed in [70]. FGAF-CDG divides the area into virtual hexagonal grid cells first, then places the cells according to their geographic location. A FL algorithm is used to choose the CH sensor in each grid cell during each sample round. In the compressive DC phase, CH readings would be delivered to the sink through a multi-hop path formed using a fuzzy-based routing mechanism.

At the WSN fusion center, a Deep Learning Based Data Mining (DDM) model is employed to accomplish energy saving and effective load balancing, according to [71]. The RNN-LSTM recurrent neural network (RNN) based Long-Short-Term Memory (LSTM) in the provided DMM model splits the network into layers and places them in the nodes. The suggested model decreases the fusion center's overhead while also reducing the quantity of data transfers. The provided RNN-LSTM model is put to the test using a variety of different hidden layer nodes and signaling intervals. Simultaneously, the amount of energy required to transfer data by the RNN-LSTM model is far less than that required to transmit actual data.

Cluster-Tree based approach (CTEEDG) of [72] enhances the throughput and lifespan of WSN. FL is used for CH selection based on locally obtained information. The tree topology from the clusters towards the BS is developed in the inter-cluster communication process, ensuring the usability of the best congestion-free route to BS.

Another DC approach in disjoint WSN using a rendezvous selection scheme with mobile edge nodes is introduced in [73]. This paper analyzed the disconnected WSNs and suggested a rendezvous selection technique, called Divide-and Rule ACO (DR-ACO), for maximum network connectivity and low propagation delay in delay-harsh applications. The route segmenting method and the candidate grouping process are introduced, which collectively enforce the principle of divide-and-rule and thus ease the ACO algorithm's mission. The powerful pheromone strength and heuristic factor are built during the transition of the ant, which collectively decrease the route length. The work presented by [74] combines rough set theory with an optimized convolutional neural network and suggests a WSN data aggregation scheme. First of all, in sink node, a feature extraction model is developed and then trained, where the rough set principle is implemented to utilize knowledge efficiently and reduce the tagged dimension.Once the cluster nodes have extracted these properties of the data from the granular deep network, CHs send them to the sink node to reduce data transfer and prolong the WSN's lifetime.

To maximize WSN lifetime and to ensure data aggregation with less energy usage for distributed WSN, the authors of [75] suggested a method integrating grid clustering and fuzzy reinforcement-learning. Initially, for cluster forming and CH collection, grid clustering is employed. In addition, depending on the parameters, such as size, algebraic connectivity and neighborhood overlap, a fuzzy rule system-based reinforcement learning method is utilized to pick the data aggregator node. Finally, the MS's hierarchical relocation is done using a fruit fly optimization algorithm within a grid-based clustered network area.

'Monkey Tree Search-based Location-Aware Smart Collector (MTS-LASC)' that utilizes fauna inspired Monkey Tree Search (MTS) behavioral model is developed in [76]. MTS-LASC is a highly dynamic and interesting phenomenon that involves dispersed smart collectors and a centralized MTS meta-heuristic engine to solve complicated and hard problems. With a client MTS module, the distributed smart collector is embedded. It can evaluate, categorize and aggregate the gathered data and disseminate them to the sink using the fuzzy inference process, whereas the centralized MTS engine uses meta-heuristic search to promote detailed understanding of the situation across multiple paths for selecting an energy-efficient path in serving critical decision-making of IoT-applications. In [77], a periodic multi-node charging and data gathering method using Mobile Device (MD) is proposed to provide a perpetual network service where the network is split into several cells and the MD passes through each cell periodically to gather data and to charge the nodes with the goal of enhancing the volume of data generated per MD's unit energy. The anchor optimization technique for optimizing the anchor point location in the cell is presented in order to realize multi-node charging. A Discrete Fireworks Algorithm (DFWA) based on population entropy is then proposed to solve the issue of mobile device route planning. In [78], a clustering based fuzzy C-Means algorithm is proposed for similarity-aware data aggregation in clustered WSNs by investigating the temporal and spatial correlation of data and the local detection of abnormal events. Then, for further outlier detection, a support degree function is defined. Finally, the aggregation is performed according to the statistical analysis of the outlier or non-outlier sensor data.

Ref. [79] proposes an Optimal Partial Aggregation (OPA) approach for balancing energy and latency utilizing several optimization techniques. The Modified multi-fruit Fly Optimization Method (MFOA) optimizes the time-varying node lifespan, while the modified Time on Task optimization algorithm minimizes the aggregation delay. The Non-dominated Sorting Gravitational Searching Algorithm is then used to calculate the routing route to the destination (NSGSA).

In [80], the nodes are arranged in a tree topology, with data aggregation taking place at intermediary nodes at tree branch junctions. To build an effective data aggregation tree, the River Formation Dynamics (RFD) method is used. The Improved RFD (IRFD) technique is then used to build a better aggregation tree. IRFD can generate more efficient aggregation trees with a reduced tree size.

In the context of assuring network transmission reliability for WSN, the challenge of energy-efficient data collecting using in-network aggregation is studied in [81]. To increase energy efficiency while ensuring request transmission reliability, a ring and fuzzy rule-based data aggregation architecture is proposed. The count of unicast packet copies is adaptively changed using FL. Based on the request transmission reliability and node energy cost imbalance, the suggested approach adaptively unicasts a variable number of aggregated packet copies in a window.

In [82] a Fuzzy-based Mobile Agent Migration (FuMAM) approach to find an appropriate mobile agent path is introduced. Three characteristics are taken into account by FuMAM: distance, remaining energy, and the count of neighbors. FuMAM improves the round-trip rate of successful mobile agents. FuMAM further extends the life of the network by choosing the node with the highest residual energy as the next hop for mobile agent migration.

To address the problem of energy consumption and packet delivery ratio in WSNs, a fuzzy-based on-demand energy efficient clustering technique is developed in [83]. It applies ACO for mobile collector movement. PSO was utilized to optimize the membership functions of the FL controllers. In terms of network longevity and packet delivery, the results have improved.

A approach to efficiently regulate the Mobile Base Station (MBS) mobility and locate the optimum destination in real time is presented in [84]. Type-2 Fuzzy Controllers control MBS movement (T2-FLCs). T2-FLCs use WSN local information to prioritize which virtual clusters should be accessed. As a result, the MBS dynamically determines which cluster is the most important to visit.

In [85], an energy-efficient Voronoi Fuzzy multi hop Clustering (V-FCM) technique for WSN data aggregation is discussed. It combines the Voronoi diagram and a modified Fuzzy C-Means algorithm. To limit the quantity of data transfers, it uses data aggregation algorithms including MAX, MIN, and AVG, which are computed in each CH. By combining Voronoi diagrams into distributed clustering approaches, V-FCM achieves its purpose. The use of Voronoi diagrams is robust, and its major advantage is that it reduces the count of Euclidean distance computations in the fuzzy clustering method and identifies the best path for data transfer. In path identification for data transfer, the Euclidean least spanning tree along with Delaunay triangulation is applied. Energy-efficient routing algorithm to prolong lifetime (ERAPL) of WSN is proposed in [86]. In this work, a data gathering sequence (DGS) is constructed to eliminate mutual transmission and loop transmission among nodes, whereby each node proportionally forwards traffic to its neighboring node. ERAPL can improve network lifetime while expending energy efficiently through the construction of DGS and selection of optimal OTPs (outgoing traffic proportion (OTP) from node i to node j) for all the nodes to distribute packets to their respective neighboring nodes.genetic algorithms. GAs with compressed chromosome coding scheme are used to find the optimal OTP matrix.

An efficient distributed intelligent data-gathering algorithm called DIDGA is proposed in [87] for the mobile collector. A mobile collector is employed to gather the sensed data from nodes by dividing the whole network into certain minimum connected dominating set (MCDS) that could minimize the number of hops in the network. A path formation optimized algorithm (PFOA) is also proposed which combines ant colony algorithm and evolutionary algorithm to satisfy the time-limit constraints.

In [88] a meta-heuristic optimization technique, Cuckoo Search (CS), is used to aggregate data in the WSN. In the proposed technique, the least energy nodes are formed as subordinate chains (or) clusters for sensing the data, and high energy nodes as Cluster Head for communicating to the Base Station (BS). The modified CS is proposed to get enhanced network performance incorporating balanced energy dissipation and this results in the formation of optimum number of clusters and minimal energy consumption. Bees algorithm is employed and modified by implementing custom crossover operation during the global search step in Intelligent DC Technique [89]. The main target of the modified Bee is to form disjoint dominating sets that work as collectors in each round. Intelligent Proficient DC Approach (IPDCA) [90] utilizes public vehicles as the mobile data collectors (D-collectors) that read (or collect) data from multiple Access Points (APs) and send them back to the central Base Station (BS). IPDCA adopts a modified Bat algorithm for path finding of D-collectors, where the Bat algorithm is extended to solve a discrete optimization problem. Fuzzy attribute-based joint integrated scheduling and tree formation (FAJIT) technique for tree formation and parent node selection using fuzzy logic in a heterogeneous WSN is proposed in [91]. In FAJIT, fuzzy logic is first applied to WSN, and then min-max normalization is used to retrieve normalized weights (membership values) for the given edges of the graph. This membership value is used to denote the degree to which an element belongs to a set. Therefore, the node with the minimum sum of all weights is considered as the parent node. A Fuzzy-based data aggregation scheme called F-LEACH is proposed in [92] with the aim to maximize the network lifetime by optimizing the aggregating node selection. Besides, in F-LEACH, the FIS membership functions are optimized, and the average of several executions are selected as the optimal parameters by modifying the network scenario. Cluster-based data aggregation method using multi-objective male lion optimization algorithm (DA-MOMLOA) for evaluating the network based on energy, delay, density and distance is presented in [93]. In DA-MOMLOA, the data aggregation method is employed with the help of cluster head wherein the data aggregated from similar clusters are forwarded to the sink node. Query ordering with data aggregation is the process of scheduling the nodes to receive the useful data from sensors. Ref. [94] proposes a novel query-based data aggregation model with the aid of intelligent techniques. The framing of the query order takes place and the frames are ranked on the basis of a multi-objective function. The newly developed multi-objective function includes Latency, Throughput, and Data freshness. Initially, the solution corresponding to query order is trained in NN using the proposed Fitness-Mated Lion Algorithm (FM-LA). The optimally generated query order from Neural Network (NN) is further given for second-level solution generation, which is again applied to FM-LA for subsequent query order optimization. Hence the two-stage optimization process with NN for query ordering is compared over the conventional methods in terms of performance measures like Latency, Throughput, and Data freshness. Ref. [95] proposed multi-weight chicken swarm based genetic algorithm for energy efficient clustering (MWCSGA) protocol to increase the energy efficiency during

the process of communication in the network. MWCSGA consists of six sections. They

are system model, chicken swarm optimization, genetic algorithm, CCSO-GA cluster head selection, multi weight clustering model, inter cluster, and intra cluster communication. In [96], Q-learning-based data-aggregation-aware energy-efficient routing algorithm

(Q-DAEER) uses reinforcement learning to maximize the rewards, defined in terms of the efficiency of the sensor type-dependent data aggregation, communication energy and node residual energy, at each sensor node to obtain an optimal path. In Q-DAEER, a data-type-dependent action selection and Q-table updating algorithm is incorporated. Adaptive routing algorithm for in-network aggregation (RINA) is proposed in [97]. RINA employs a reinforcement learning method called Q-learning to build a routing tree based on minimal information such as residual energy, distance between nodes and link strength. In addition, RINA finds the aggregation points in the routing structure to maximize the number of overlapping routes in order to increase the aggregation ratio.

A framework for long-term monitoring applications called spatiotemporal approximate DC (STAC) is proposed in [98]. This framework depends on WSNs or IoT and utilizes data prediction, selecting work nodes, and Q-learning. In STAC, the spatio-temporal correlation among nodes is considered. First, data prediction by selecting work nodes in each work cycle is performed. Second, takes advantage of temporal redundancy and employ Q-learning to adjust the sampling interval of nodes. Finally, the BS predicts the whole data with error tolerance.

Table 2. Summary of AI based solutions to WSN DC, aggregation and dissemination issues surveyed in Section 4.

AI Paradigms	Algorithm	Ref.	Optimization Criteria	Simulation/ Real-Deployment	Centralized/ Distributed	Mobility	Performance Parameters
Swarm Intelligence	DAACA	[44]	Minimizing energy consumption	Java Simulation	Distributed	Static	Average energy cost, Energy difference, WSN lifetime, One hop transmission success ratio and time complexity.
Swarm Intelligence	ACOPSO	[55]	Improving inter-cluster data aggregation	MATLAB simulation	Distributed	Static	Stability period, Network lifetime, Average remaining energy, and Throughput.
Swarm Intelligence	ABC	[47]	Optimizing energy consumption	Random network deployment	Distributed	Static	Average path length
Swarm Intelligence	AMIDT	[46]	Reducing minimum-power multiresolution data dissemination problem	GloMosim simulator with PARSEC compiler	Distributed	Static	Energy cost for path finding, and Total power consumption.
Swarm Intelligence	ALOC	[54]	Improving lifetime and throughput of a network	MATLAB simulation	Distributed	Static (MS)	Dead nodes, Alive nodes, Residual energy, Throughput, Lifetime, Number of individual nodes, Number of CHs, and DC tour length.
Swarm Intelligence	HPGSO-ADC	[51]	Improving quality of service	Real network deployment	Distributed	Mobile	Network lifetime,Data packets delivered to BS, Energy consumption, and Alive nodes.
Swarm Intelligence	SDMA-PSO	[56]	Improving efficiency and reduces buffer overflow	NS2 simulator	Distributed	Static (MS)	Delay, Throughput, Average energy consumed, and Network lifetime.
Swarm Intelligence	CC-MRT, CC-TRT, CC-WMRT	[66]	Improving network lifetime	Simulation	Distributed	Static	Transmissions count, leftover energy, energy usage variance, and WSN lifetime
Swarm Intelligence	DFWA	[77]	Maximizing the amount of data collected	Simulation	Distributed	mobile	Cell numbers and Sojourn time.

AI Paradigms	Algorithm	Ref.	Optimization Criteria	Simulation/ Real-Deployment	Centralized/ Distributed	Mobility	Performance Parameters
Swarm Intelligence	PSO	[52]	Parameter optimization	Matlab simulation	Distributed	Static	Total error probability and Time consumed for convergence.
Swarm Intelligence	Firefly Optimization	[57]	Maximize the network lifetime	MATLAB simulation	Distributed	Static (MS)	Network lifetime, Average energy consumption, and First and last dead nodes.
Swarm Intelligence	Shuffled Frog	[49]	Reducing energy consumption	MATLAB simulation	Centralized	Static	Average end to end delay, Lifetime, Average packet loss rate, Count of clusters, and Residual Energy.
Swarm Intelligence	CoDA	[50]	Minimum power consumption	NS2 simulation	Centralized	Static	Synchronization error, Synchronization overhead Energy consumption, and Average path length.
Swarm Intelligence	PSO	[53]	Optimal service compositions	Java simulation	Centralized	Static	Energy consumption.
Swarm Intelligence	DR-ACO	[73]	Full connectivity and minimize the data transmission delay	Matlab simulation	Distributed	Static with mobile edge nodes	Path length, Distance Limit , and Connectivity Degree.
Evaluation Computation	MTS-LASC	[76]	Low delivery latency, Reasonable and balanced Energy usage	Matlab simulation	Distributed	Static	Delay, Routing overhead, Packet delivery ratio.
Swarm Intelligence	CSO	[60]	Optimize CS matrix	Matlab simulation	Centralized	Static	Average energy consumption, Network lifetime, Number of data transmissions, and Average Normalized MSE.
Trajectory based	SATC	[61]	Minimizing Aggregation Time	Simulation	Centralized	Static	Average Latency and Average Normalized Latency.
Evolutionary Computation	ETDMA-GA	[62]	Minimizing the data delivery time	Simulation	Centralized	Static	Schedule Length, Duty Cycle, Average Latency, and Average Normalized Latency.
Physical Computation	IWDs	[45]	Energy conservation	Simulation	Distributed	Static	Total energy consumption, Network lifetime, and Time Complexity.
FL	DFHFE	[63]	Energy consumption	MATLAB simulation	Distributed	Static	Overall network energy consumption , WSN lifetime, Alive nodes.
FL	TTDFP	[58]	Efficiency of data aggregation operations	MATLAB simulation	Centralized	Static	First Node Die, Half Nodes Die, and Total Remaining Energy.
FL	FGAF-CDG	[70]	Improve energy consumption and load balancing	MATLAB simulation	Distributed	Static	Energy consumption and Network lifetime.
FL	CTEEDG	[72]	Increase the lifetime and throughput	NS2 simulation	Distributed	Static	Average end-to-end delay, Control overhead, Throughput, Alive nodes.
A Hybrid of FL and SI	MDF-FBCHS	[59]	Maximizing network lifetime	SimpleIoTSimulator	Distributed	Static	Alive nodes, WSN lifetime, end-to-end delay, Energy usage, Average delay, Throughput, and Packet delivery ratio.
A Hybrid of FL and SI	FCOABC	[64]	Increasing network lifespan and energy efficiency	NS2 and MATLAB simulation	Distributed	Static	Data latency, Delivery ratio, Energy consumption, Network lifetime, and Complexity analysis.
DL	DNN	[48]	Data mining	MATLAB simulation	Distributed	Static	Fault detection rate, Energy usage rate.
A Hybrid of ACO and Game theory	GTAC-DG	[65]	Increasing network lifespan and energy efficiency	MATLAB simulation	Distributed	Static with MS	Energy consumption, Network lifetime, Total travel time and Overlapping area.

Table 2. Cont.

AI Paradigms	Algorithm	Ref.	Optimization Criteria	Simulation/ Real-Deployment	Centralized/ Distributed	Mobility	Performance Parameters
A Hybrid of FL, Reinforcement- Learning and Fruit Fly optimisation	FRS-RL	[75]	Increasing network lifespan and energy efficiency	MATLAB simulation	Distributed	Static	Packet delivery ratio, Latency, Packet loss ratio, Throughput, Energy consumption, WSN lifetime.
A Hybrid of FL and Harris Hawks Optimization (HHO) algorithm	RDDI	[67]	Energy efficiency, Reliable and secure data aggregation	MATLAB simulation	Distributed	Static	Energy consumption, Total distance, Average end-to-end delay, Reliability, and Overhead analysis.
A Hybrid of PSO and Neural Network	PSO-ELM	[69]	Prolong network lifetime	MATLAB simulation	Distributed	Static	Energy Consumption, Node Survival, Network Load Balancing, Network Connectivity, and Reliability.
Neural Network	Convolutional Neural Network	[74]	Accuracy of data aggregation	Simulation based Keras framework	Distributed	Static	Data fusion accuracy, Network lifetime, and Energy consumption.
DL	RNN-LSTM	[71]	Energy efficiency and optimal load balancing	MATLAB simulation	Distributed	Static	Signaling overhead, Average throughput, and Average delay.
FL	Fuzzy C-Means	[78]	Optimize clustering based on data similarity	TOSSIM	Distributed	Static	False alarm rate, Data outlier detection accuracy, and Relative recovery error.
A Hybrid of Physical and Nature Inspired computation	MFOA and NSGSA	[79]	Maximizing network lifetime and efficient routing path	NS2	Distributed	Static	Delay, Energy usage, WSN lifetime, Delivery ratio, Throughput.
Physical Computation	RFD and IRFD	[80]	Optimize data aggregation tree	Simulation	Distributed	Static	Energy consumption, and Network lifetime.
FL	Ring based and FL	[81]	Optimize in-network data aggregation	MATLAB simulation	Distributed	Static	Reliability, Energy cost, and Network lifetime.
FL	FuMAM	[82]	Optimal itinerary for mobile agent	MATLAB simulation	Distributed	Static with mobile agent	Round-trip rate, WSN lifetime, Energy usage, Task duration, Energy distribution.
A Hybrid of Fuzzy and SI	Fuzzy and PSO	[83]	Optimize membership functions of the FL and mobile collector movement	MATLAB simulation	Distributed	Static with mobile data collector	Packet delivery ratio, Clustering overhead, and Network lifetime.
FL	T2-FLCs	[84]	Best location and DC time	MATLAB simulation	Distributed	Static with MS	Average energy consumption, End-to-end delay, Data loss ratio, and Unsuccessful transmission.
FL	V-FCM	[85]	Optimize clustering along with data aggregation	MATLAB simulation	Centralized	Static	Running time, Size of received data, and Alive nodes.
Evolutionary computation	ERAPL	[86]	optimize network lifetime.	simulation	Distributed	Static	Network lifetime.
Swarm Intelligence	PFOA	[87]	optimize path formation.	NS2 simulation	Distributed	Static with mobile collector	average hop counts, average data-gathering time and Network lifetime.
Nature inspired	CUCKOO	[88]	optimize network lifetime	Matlab simulation	Distributed	Static	network Lifetime and energy consumption.
Swarm Intelligence	IDCT based Bees algorithm	[89]	construct disjoint dominating sets	MATLAB simulation	Centralized	Static	Network lifetime and Average consumption of energy.
Nature inspired	IPDCA based BAT algorithm	[90]	Path finding of Data collectors	MATLAB simulation	Centralized	Static with mobile data collector	AP Count,D-collector count,Cost, Throughput,Area size,APs Load, and D-collector Capacity.

Table 2. Cont.

AI Paradigms	Algorithm	Ref.	Optimization Criteria	Simulation/ Real-Deployment	Centralized/ Distributed	Mobility	Performance Parameters
Q-learning	Q-DAEER	[96]	Maximize the lifetime and minimize energy consumption of the network	MATLAB simulation	Centralized	Static	Network-level energy consumption, number of dead nodes,network lifetime, average hop count and decrease in data size.
FL	FAJIT	[91]	parent node selection and improving energy efficiency in WSN.	Simulation	Distributed	Static	Schedule length ,Average aggregation factor,Control overhead, Average energy consumption , and Number of transmission slots.
FL	F-LEACH	[92]	maximize the network lifetime.	Simulation	Distributed	Static	Average residual energy and percentage of dead nodes.
Nature inspired	DA- MOMLOA	[93]	reducing the cost of transmission	Simulation	Distributed	Static	Network efficiency, Scala- bility,Throughput,Data delivery, Packet drop.
A Hybrid of Nature inspired and Artificial neural network	FM-LA with NN	[94]	Optimal data aggregation	MATLAB simulation	Centralized	Static	latency, throughput, and data freshness.
A Hybrid of Nature inspired and Evolution Computation	MWCSGA	[95]	improve the life span of the network, energy efficiency enhancement	NS2 simulation	Centralized	Static	energy efficiency, energy consumption, end to end delay, packet drop, and throughput and packet delivery ratio.
Q-learning	RINA	[97]	Increase aggregation ratio	OMNet++ simulation	Distributed	Static	Network scalability, Events, Communication range, Communication failure, Aggregation ratio, Network load.
Q-learning	STAC	[98]	Reduce the data redundancy	MATLAB simulation	Distributed	Static	network lifetime and error of predicted data.

Table 2. Cont.

4.2. Open Research Issues and Challenges

In future, researchers can conduct further studies regarding outlier identification and analysis in terms of multi-dimension, detection mode, architectural structure, correlation extraction, etc. The approaches, e.g., IRFD method as discussed in [80] can be further evaluated with respect to similar algorithms like Intelligent Water Drop algorithm to identify the change in performance and can be studied for any enhancement.

For fuzzy based schemes, the influence of membership functions and fuzzy output space on performance are significant areas for further research. For example, FuMAM [82] can be evaluated further by considering different input parameters such as remaining energy of the candidate's neighbors to Fuzzy Logic System (FLS). This addition can improve energy balance across nodes for each estimated path. Increasing the number of input parameters, on the other hand, will increase the complexity of the rule base. As a result, an approach like hierarchical fuzzy system can be tested for use in resizing the rule base. Full network coverage is indeed a metric that can be used while determining the CHs that perform the aggregation task. In addition, the techniques devised for proactive networks can be extended for use in a reactive network to satisfy intrusion detection challenges that can occur during the DC process.

The researchers can also be encouraged to test the working of proposed schemes in a real-world context to ensure that the techniques are viable. In the future, researchers can plan to broaden their existing research to include a more generic network model with multiple mobile robots using different optimization strategies. An in-depth analysis should be performed to take into account both mobile robot collaboration and travel path planning in complex and challenging environments. Minimum-power Multiresolution Data Dissemination problem can be further examined using different meta-heuristics solutions to evaluate the performance results. We hope this can help in realizing the best possible solution to this challenge.

Aside from ensuring secure and reliable information exchange between the CH and BS, researchers can also experiment with improving WSN QoS parameters (such as availability, reliability, and serviceability) by combining various optimization algorithms to achieve improved data dissemination and gathering efficiency. The researchers can investigate about the integration of different powerful data fusion algorithms into other system research areas, such as big data systems, key-value stores, data compression, and blockchains. The optimization of different algorithm dependent parameters can be studied further to explore the behavior and to test the performance under WSNs coupled with different node mobility approaches. The use of hybrid compressive sensing can be encouraged with different optimization strategies. Performing real-time implementation for related applications are of great significance as efficient implementation of general real-time applications remains by far an open challenge.

Additional research is necessary to enhance sample data noise filtering and data mining using deeper DNN layers. Moreover, improved and flexible protocols should be developed to deeply evaluate the traffic on different layers using complex functions. Further studies can be performed to optimize the algorithm using different parameter choices to adapt them for use in large scale heterogeneous networks, and simplifying the data fusion task.

For improved outcomes, a thorough examination of the sophisticated computations of intelligent optimization algorithms and hyper parameter tweaking approaches is required. The addition of new nodes following the node's death preserves the network's interaction capabilities. The technique of dynamic insertion of new nodes requires further investigation to study the impact of DC and fusion performance of an algorithm. These are key problems to examine in the future in order to improve the data gathering, aggregation, and dissemination performance of large-scale WSNs.

5. Conclusions and Future Directions

Nowadays, we are witnessing a growth in the number of AI -based systems and solutions which facilitate the optimization of services in the field of WSN. The combination of both AI methods and WSNs have now become a reality, offering benefits to the area of Internet of Things, and allow systems to learn and to monitor activities and support the decision-making process. In this paper, we have provided a review of DC, Aggregation and Dissemination challenges in WSNs. Various AI methods in WSNs are briefly introduced along with their classifications. AI techniques used by researchers to address the DC, Aggregation and Dissemination challenges in WSNs are briefly explained for the span of 2010 to 2021. AI-based solutions to various problems in WSNs have been discussed and summarized.

From the findings depicted in Figure 5 and Table 2, we can notice that Swarm intelligence is adopted as the best suited technique or acceptable method by the research community for solving DC, aggregation and dissemination challenges in WSNs. 36% of articles applied swarm intelligence while 20%, 18%, and 5% for FL, Hybrid, Reinforcement learning respectively. It can be noted that SI methods are popularly applied in WSN.

Some AI algorithms have been used frequently to solve WSN challenges. GA, PSO, ACO, ABC, BA, DL and fuzzy have been used more than other algorithms by the research community to deal with WSN challenges. These algorithms are selected based on problem nature or algorithm features. If problem parameters have fuzzy feature and membership function, FL method is suitable to deal with this problem. When problem size is too vast, and needs to be adapted to a particular case and with respect to time, we need to solve it using DL.

GA is very efficient and stable in exploring the search space for global optimal solutions. It helps to solve general, unconstrained and bound constrained optimization problems, and it performs very well for large scale optimization problems. Both con-



tinuous and discrete parameters can be used [99–101]. So researchers use GA for these characteristics that help to solve a wide variety of WSN challenges.

Figure 5. AI methods with respect to various challenges in WSN.

From Table 2, we can notice that PSO is used commonly in WSN. This is due to the fact that PSO has a number of distinct advantages over other optimization strategies. The PSO algorithm is a derivative-free method. It can be used in conjunction with other optimization strategies. It is less sensitive to the objective function's nature and can handle stochastic objective functions. PSO has a small number of parameters, can execute parallel computations, converges quickly, and is simple to implement [101].

Working of ABC is focussed on the behavior of bees in visiting and selecting best source of food (flowers). So, it works well in selection problems such as choosing best CHs in WSN. ABC is easy to use, available for hybridization with other algorithms [101].

As we have described above, population based meta-heuristic algorithms are used extensively because of their efficiency in providing best results. Their individuals don't use any symbolic reasoning and don't maintain a plan about the future, both of which are computationally expensive. Moreover, these techniques involve only limited memory requirements as they don't need to keep a lot of previous information. It can be robust, and maintain good quality performance in rapidly changing and diverse environments.

As we notice from Figure 5 and Table 2, some AI methods such as RL. The deep reinforcement learning is also used rarely in WSN. This is because DRL faces main challenges like non-stationary environment, partial observability of the environment, continuous action spaces and is computationally expensive [102].

For enhancement of WSN, new AI algorithms as well as different strategies to embed these algorithms in WSNs have to be encouraged in future. Currently, the most of the solutions presented here used AI to solve specific challenges in some areas. Most problems stem from incompatibility between layers and high human interaction. Self-Adaptivity is required for setting and adjustment of solutions. Hybrid approaches that optimize the resource utilization in WSN need to be developed [27]. Rather than specialized solutions, learning platforms and prototypes are required.

In future, further research to devise efficient distributed data mining solution for WSN with improvements in noisy data filtering process have to be encouraged. The layered structure and other typical features of DNN makes it a favorable option for application in such scenarios. The main challenges include training of distributed multi layered DNN and better tradeoff between power consumption during processing and transmission. A

further interesting area of study for AI methods in WSNs is known to be parameter learning

and optimization. Many efforts have been made to fix the interference classification challenge in WSNs. The steadily increasing utilization of license-free frequency bands make proper wireless interference identification and reliable management a crucial challenge. Interference mitigation techniques, deployment planning and other potential applications opens up numerous possibilities for future research. DL classification model is a choice.

Traffic flow control in WSN is a potential future direction. Mobility is another important open research topic that can be supported by head nodes, normal nodes, and BS. The most significant issues to be addressed are changes in WSN architecture and control packet overhead caused by node mobility. Only a few researches have been conducted to improve the intelligence of mobile agent for better route planning and data gathering. To accommodate the node changes in a dynamic WSN, an integration of DL and RL using multiple mobile agents is a desirable solution. The use of RL can make mobile agents more intelligent in making decisions to perform best action according to the present situation. Moreover, determining optimal number of mobile agents and their route by evaluating multiple parameters that could provide efficient results require further investigation in future. Moreover, most of the current algorithms didn't properly consider the security and privacy of the data in multiple mobile agent scenarios. The utilization of compression techniques to compress the data collected by each mobile agent also require more research attention. It can be expected that the future researches will give better consideration to these issues. Future development of more robust and efficient mobile agent learning methods to solve the real-world challenges are necessary.

The application of AI optimization methods to overcome the challenges of MWSNs is a potential future direction. Some research challenges still remain relatively unresolved which include transmission delay, balancing the power consumption, reliability and safety of MWSNs. A combination of SI with other optimization techniques have to be encouraged in future. The cross-layer optimization model challenges must to be treated in a better way during the optimization of MWSNs. The results from the study of human-related biological features can be incorporated in future for further improvement of solutions to such problems. Distributed and real time application of algorithms in light weight form could be a question for further studies in order to solve the challenges of dynamic MWSNs in future.

As the study shows, most of the existing solutions based on AI are only simulation based. The AI techniques should be implemented and analyzed in a real-time environment [26,28]. This should be encouraged in future. More investigations are required to show how AI methods can be adapted. Cross-layer approaches using AI methods are rarely applied to challenges and are still a vital open research area. Moreover, hybrid AI methods are less applied and need to be discussed deeply. It can be expected that the future research will likely consider heterogeneity, dynamic environments and varying communication constraints during algorithm design. The application of AI techniques brings the WSN cognitive towards managing and overcoming the challenges which arise during operation. We hope that the concepts provided in this paper direct the researchers for the use of AI in solving the challenging WSN issues by making the nodes more intelligent.

Author Contributions: All authors contributed equally to this paper, where W.O., A.M.K., A.S., A.A.E.-S. and A.I.A. participated in sorting the experiments, discussed and analyzed the results, performed the experiments and analyzed the results, wrote the paper, discussed the result, and review/edited the manuscript. All authors have read and agreed to this version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The study did not report any data.

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

References

- Bellavista, P.; Cardone, G.; Corradi, A.; Foschini, L. Convergence of MANET and WSN in IoT urban scenarios. *IEEE Sens. J.* 2013, 13, 3558–3567. [CrossRef]
- Osamy, W.; Khedr, A.M.; Aziz, A.; El-Sawy, A.A. Cluster-tree routing based entropy scheme for data gathering in wireless sensor networks. *IEEE Access* 2018, 6, 77372–77387. [CrossRef]
- 3. Omar, D.M.; Khedr, A.M.; Agrawal, D.P. Optimized clustering protocol for balancing energy in wireless sensor networks. *Int. J. Commun. Netw. Inf. Secur.* 2017, 9, 367–375.
- 4. Jabbar, W.A.; Saad, W.K.; Ismail, M. MEQSA-OLSRv2: A multicriteria-based hybrid multipath protocol for energy-efficient and QoS-aware data routing in MANET-WSN convergence scenarios of IoT. *IEEE Access* **2018**, *6*, 76546–76572. [CrossRef]
- Bradai, A.; Benslimane, A.; Singh, K.D. Dynamic anchor points selection for mobility management in Software Defined Networks. J. Netw. Comput. Appl. 2015, 57, 1–11. [CrossRef]
- Amri, S.; Khelifi, F.; Bradai, A.; Rachedi, A.; Kaddachi, M.L.; Atri, M. A new fuzzy logic based node localization mechanism for wireless sensor networks. *Future Gener. Comput. Syst.* 2019, 93, 799–813. [CrossRef]
- Khalifeh, A.; Darabkh, K.A.; Khasawneh, A.M.; Alqaisieh, I.; Salameh, M.; AlAbdala, A.; Alrubaye, S.; Alassaf, A.; Al-HajAli, S.; Al-Wardat, R.; et al. Wireless Sensor Networks for Smart Cities: Network Design, Implementation and Performance Evaluation. *Electronics* 2021, 10, 218. [CrossRef]
- Khedr, A.M.; Omar, D.M. SEP-CS: Effective Routing Protocol for Heterogeneous Wireless Sensor Networks. Ad Hoc Sens. Wirel. Netw. 2015, 26, 211–234.
- 9. Khedr, A.M.; Bhatnagar, R. Agents for integrating distributed data for complex computations. Comput. Inform. 2007, 26, 149–170.
- 10. Khedr, A.M. Learning k-nearest neighbors classifier from distributed data. Comput. Inform. 2008, 27, 355–376.
- 11. Wang, F.; Liu, J. Networked wireless sensor data collection: Issues, challenges, and approaches. *IEEE Commun. Surv. Tutor.* **2010**, 13, 673–687. [CrossRef]
- 12. Randhawa, S.; Jain, S. Data aggregation in wireless sensor networks: Previous research, current status and future directions. *Wirel. Pers. Commun.* **2017**, *97*, 3355–3425. [CrossRef]
- 13. Hamida, E.B.; Chelius, G. Strategies for data dissemination to mobile sinks in wireless sensor networks. *IEEE Wirel. Commun.* **2008**, *15*, 31–37. [CrossRef]
- 14. Bello, O.; Holzmann, J.; Yaqoob, T.; Teodoriu, C. Application of artificial intelligence methods in drilling system design and operations: A review of the state of the art. *J. Artif. Intell. Soft Comput. Res.* **2015**, *5*, 121–139. [CrossRef]
- 15. Luke, S. *Essentials of Metaheuristics*, 2nd ed.; Lulu Enterprises: Wuhan, China, 2013. Available online: http://cs.gmu.edu/~sean/book/metaheuristics/ (accessed on 1 August 2021).
- 16. Yang, X.S. Nature-Inspired Metaheuristic Algorithms; Luniver Press: Bristol, UK, 2010.
- 17. Xing, B.; Gao, W.J. Fruit fly optimization algorithm. In *Innovative Computational Intelligence: A Rough Guide to 134 Clever Algorithms;* Springer: Berlin/Heidelberg, Germany, 2014; pp. 167–170.
- 18. Brownlee, J. Clever Algorithms: Nature-Inspired Programming Recipes; Lulu Enterprises: Wuhan, China 2011.
- 19. Jagtap, V. Survey of different swarm intelligence algorithms. Int. J. Adv. Eng. Res. Dev. 2014, 1, 12.
- 20. Dorri, A.; Kanhere, S.S.; Jurdak, R. Multi-agent systems: A survey. IEEE Access 2018, 6, 28573–28593. [CrossRef]
- 21. Chen, S.H.; Jakeman, A.J.; Norton, J.P. Artificial intelligence techniques: An introduction to their use for modelling environmental systems. *Math. Comput. Simul.* **2008**, *78*, 379–400. [CrossRef]
- 22. Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, M.S.; Hasan, M.; Van Essen, B.C.; Awwal, A.A.; Asari, V.K. A state-of-the-art survey on deep learning theory and architectures. *Electronics* **2019**, *8*, 292. [CrossRef]
- 23. Chang, C.W.; Lee, H.W.; Liu, C.H. A review of artificial intelligence algorithms used for smart machine tools. *Inventions* **2018**, 3, 41. [CrossRef]
- 24. Boucher, P. *How Artificial Intelligence Works*; Scientific Foresight Unit; EPRS | European Parliamentary Research Service, European Union: 2019. Available online: https://www.europarl.europa.eu/at-your-service/files/be-heard/religious-and-non-confessional-dialogue/events/en-20190319-how-artificial-intelligence-works.pdf (accessed on 23 November 2021).
- 25. Kulkarni, R.V.; Forster, A.; Venayagamoorthy, G.K. Computational intelligence in wireless sensor networks: A survey. *IEEE Commun. Surv. Tutor.* **2010**, *13*, 68–96. [CrossRef]
- 26. Sarobin, M.V.R.; Ganesan, R. Swarm intelligence in wireless sensor networks: A survey. Int. J. Pure Appl. Math. 2015, 101, 773–807.
- 27. Jabbar, S.; Iram, R.; Minhas, A.A.; Shafi, I.; Khalid, S.; Ahmad, M. Intelligent optimization of wireless sensor networks through bio-inspired computing: Survey and future directions. *Int. J. Distrib. Sens. Netw.* **2013**, *9*, 421084. [CrossRef]
- Montoya, A.; Restrepo, D.C.; Ovalle, D.A. Artificial intelligence for wireless sensor networks enhancement. In Smart Wireless Sensor Networks; IntechOpen: London, UK, 2010; pp. 73–81.
- 29. Yu, Y. A survey of intelligent approaches in wireless sensor networks for efficient energy consumption. In Proceedings of the International Conference on Software and Computer Applications (ICSCA'12), Singpore, 9–10 June 2012.
- Maksimović, M.; Vujović, V.; Milošević, V. Fuzzy logic and Wireless Sensor Networks—A survey. J. Intell. Fuzzy Syst. 2014, 27, 877–890. [CrossRef]

- 31. Dehkordi, S.A.; Farajzadeh, K.; Rezazadeh, J.; Farahbakhsh, R.; Sandrasegaran, K.; Dehkordi, M.A. A survey on data aggregation techniques in IoT sensor networks. *Wirel. Netw.* **2020**, *26*, 1243–1263. [CrossRef]
- 32. Maraiya, K.; Kant, K.; Gupta, N. Wireless sensor network: A review on data aggregation. Int. J. Sci. Eng. Res. 2011, 2, 1-6.
- 33. Dhasian, H.R.; Balasubramanian, P. Survey of data aggregation techniques using soft computing in wireless sensor networks. *IET Inf. Secur.* **2013**, *7*, 336–342. [CrossRef]
- Tsai, C.W.; Tsai, P.W.; Pan, J.S.; Chao, H.C. Metaheuristics for the deployment problem of WSN: A review. *Microprocess. Microsyst.* 2015, 39, 1305–1317. [CrossRef]
- 35. Tsai, C.W.; Hong, T.P.; Shiu, G.N. Metaheuristics for the Lifetime of WSN: A Review. IEEE Sens. J. 2016, 16, 2812–2831. [CrossRef]
- 36. Fei, Z.; Li, B.; Yang, S.; Xing, C.; Chen, H.; Hanzo, L. A survey of multi-objective optimization in wireless sensor networks: Metrics, algorithms, and open problems. *IEEE Commun. Surv. Tutor.* **2016**, *19*, 550–586. [CrossRef]
- Chen, X.; Yu, L.; Wang, T.; Liu, A.; Wu, X.; Zhang, B.; Lv, Z.; Sun, Z. Artificial intelligence-empowered path selection: A survey of ant colony optimization for static and mobile sensor networks. *IEEE Access* 2020, *8*, 71497–71511. [CrossRef]
- 38. Guidi, B.; Ricci, L. Aggregation techniques for the internet of things: An overview. In *The Internet of Things for Smart Urban Ecosystems*; Springer: Cham, Switzerland, 2019; pp. 151–176. [CrossRef]
- Ambigavathi, M.; Sridharan, D. Energy-aware data aggregation techniques in wireless sensor network. In Advances in Power Systems and Energy Management; Springer: Singapore, 2018; pp. 165–173. [CrossRef]
- 40. Saeedi, I.D.I.; Al-Qurabat, A.K.M. A Systematic Review of Data Aggregation Techniques in Wireless Sensor Networks. J. Phys. Conf. Ser. 2021, 1818, 012194. [CrossRef]
- Kiruthiga, T.; Shanmugasundaram, N. In-network Data Aggregation Techniques for Wireless Sensor Networks: A Survey. In Computer Networks, Big Data and IoT; Springer: Berlin/Heidelberg, Germany, 2021; pp. 887–905.
- 42. Labib Borham, M.; Khoriba, G.; Mostafa, M.S. Data collection protocols for wireless sensor networks. *Int. J. Electr. Comput. Eng. Syst.* 2021, 12, 207–214. [CrossRef]
- 43. Faris, K.; Asgarnezhad, R.; Fathal, Z.; Fleeh, A.; Mohammed, A. Survey of Wireless Sensor Network based on Power Consumption and Data Aggregation. *Int. J. Progress. Res. Sci. Eng.* **2021**, *2*, 22–32.
- Lin, C.; Wu, G.; Xia, F.; Li, M.; Yao, L.; Pei, Z. Energy efficient ant colony algorithms for data aggregation in wireless sensor networks. J. Comput. Syst. Sci. 2012, 78, 1686–1702. [CrossRef]
- 45. Hoang, D.C.; Kumar, R.; Panda, S.K. Optimal data aggregation tree in wireless sensor networks based on intelligent water drops algorithm. *IET Wirel. Sens. Syst.* 2012, 2, 282–292. [CrossRef]
- 46. Woungang, I.; Dhurandher, S.K.; Agnani, L.; Mahendru, A.; Anpalagan, A. An ant-swarm inspired dynamic multiresolution data dissemination protocol for wireless sensor networks. *J. Supercomput.* **2013**, *65*, 524–542. [CrossRef]
- 47. Chang, W.L.; Zeng, D.; Chen, R.C.; Guo, S. An artificial bee colony algorithm for data collection path planning in sparse wireless sensor networks. *Int. J. Mach. Learn. Cybern.* **2015**, *6*, 375–383. [CrossRef]
- 48. Li, C.; Xie, X.; Huang, Y.; Wang, H.; Niu, C. Distributed data mining based on deep neural network for wireless sensor network. *Int. J. Distrib. Sens. Netw.* **2015**, *11*, 157453. [CrossRef]
- Abirami, T.; Anandamurugan, S. Data aggregation in wireless sensor network using shuffled frog algorithm. *Wirel. Pers. Commun.* 2016, 90, 537–549. [CrossRef]
- 50. Tang, C.; Yang, N. CoDA: Collaborative Data Aggregation in Emerging Sensor Networks Using Bio-Level Voronoi Diagrams. Sensors 2016, 16, 1235. [CrossRef]
- 51. Burugari, V.K.; Periasamy, P. Multi QoS constrained data sharing using hybridized pareto-glowworm swarm optimization. *Clust. Comput.* **2019**, *22*, 9727–9735. [CrossRef]
- 52. Jiang, S.; Zhao, Z.; Mou, S.; Wu, Z.; Luo, Y. Linear decision fusion under the control of constrained PSO for WSNs. *Int. J. Distrib. Sens. Netw.* **2012**, *8*, 871596. [CrossRef]
- 53. Zhou, Z.; Xu, J.; Zhang, Z.; Lei, F.; Fang, W. Energy-efficient optimization for concurrent compositions of WSN services. *IEEE Access* 2017, *5*, 19994–20008. [CrossRef]
- 54. Yogarajan, G.; Revathi, T. Improved cluster based data gathering using ant lion optimization in wireless sensor networks. *Wirel. Pers. Commun.* **2018**, *98*, 2711–2731. [CrossRef]
- 55. Kaur, S.; Mahajan, R. Hybrid meta-heuristic optimization based energy efficient protocol for wireless sensor networks. *Egypt. Inform. J.* **2018**, *19*, 145–150. [CrossRef]
- 56. Vijayalakshmi, K.; Manickam, J.M.L. A cluster based mobile data gathering using SDMA and PSO techniques in WSN. *Clust. Comput.* **2019**, *25*, 12727–12736. [CrossRef]
- Krishnan, M.; Yun, S.; Jung, Y.M. Improved clustering with firefly-optimization-based mobile data collector for wireless sensor networks. AEU-Int. J. Electron. Commun. 2018, 97, 242–251. [CrossRef]
- Sert, S.A.; Alchihabi, A.; Yazici, A. A Two-Tier Distributed Fuzzy Logic Based Protocol for Efficient Data Aggregation in Multihop Wireless Sensor Networks. *IEEE Trans. Fuzzy Syst.* 2018, 26, 3615–3629. [CrossRef]
- Kumar, S.; Lal, N.; Chaurasiya, V.K. A forwarding strategy based on ANFIS in internet-of-things-oriented wireless sensor network (WSN) using a novel fuzzy-based cluster head protocol. *Ann. Telecommun.* 2018, 73, 627–638. [CrossRef]
- Aziz, A.; Singh, K.; Osamy, W.; Khedr, A.M. Effective algorithm for optimizing compressive sensing in IoT and periodic monitoring applications. J. Netw. Comput. Appl. 2019, 126, 12–28. [CrossRef]

- Osamy, W.; El-sawy, A.A.; Khedr, A.M. SATC: A Simulated Annealing Based Tree Construction and Scheduling Algorithm for Minimizing Aggregation Time in Wireless Sensor Networks. Wirel. Pers. Commun. 2019, 108, 921–938. [CrossRef]
- Osamy, W.; El-Sawy, A.A.; Khedr, A.M. Effective TDMA scheduling for tree-based data collection using genetic algorithm in wireless sensor networks. *Peer-to-Peer Netw. Appl.* 2020, 13, 796–815. [CrossRef]
- Wang, J.; Tawose, O.T.; Jiang, L.; Zhao, D. A New Data Fusion Algorithm for Wireless Sensor Networks Inspired by Hesitant Fuzzy Entropy. Sensors 2019, 19, 784. [CrossRef] [PubMed]
- 64. Kalaikumar, K.; Baburaj, E. Fuzzy enabled congestion control by cross layer protocol utilizing OABC in WSN: Combining MAC, routing, non-similar clustering and efficient data delivery. *Wirel. Netw.* **2020**, *26*, 1085–1103. [CrossRef]
- 65. Raj, P.V.P.; Khedr, A.M.; Aghbari, Z.A. Data gathering via mobile sink in WSNs using game theory and enhanced ant colony optimization. *Wirel. Netw.* **2020**, *26*, 2983–2998. [CrossRef]
- 66. Sadeghi, F.; Avokh, A. Load-balanced data gathering in Internet of Things using an energy-aware cuckoo-search algorithm. *Int. J. Commun. Syst.* **2020**, *33*, e4385. [CrossRef]
- 67. Seyfollahi, A.; Ghaffari, A. Reliable data dissemination for the Internet of Things using Harris hawks optimization. *Peer-to-Peer Netw. Appl.* **2020**, *13*, 1886–1902. [CrossRef]
- 68. Sarode, P.; Reshmi, T.R. Optimized Query Ordering Data Aggregation Model Using Neural Networks and Group Search Optimization in Wireless Sensor Network. *Ad Hoc Sens. Wirel. Netw.* **2020**, *46*, 189–214.
- Cao, L.; Cai, Y.; Yue, Y. Data Fusion Algorithm for Heterogeneous Wireless Sensor Networks Based on Extreme Learning Machine Optimized by Particle Swarm Optimization. *J. Sens.* 2020, 2020, 2549324. [CrossRef]
- Ghaderi, M.R.; Vakili, V.T.; Sheikhan, M. FGAF-CDG: Fuzzy geographic routing protocol based on compressive data gathering in wireless sensor networks. J. Ambient Intell. Humaniz. Comput. 2020, 11, 2567–2589. [CrossRef]
- Mohanty, S.N.; Lydia, E.L.; Elhoseny, M.; Al Otaibi, M.M.G.; Shankar, K. Deep learning with LSTM based distributed data mining model for energy efficient wireless sensor networks. *Phys. Commun.* 2020, 40, 101097. [CrossRef]
- 72. Karunanithy, K.; Velusamy, B. Cluster-tree based energy efficient data gathering protocol for industrial automation using WSNs and IoT. J. Ind. Inf. Integr. 2020, 19, 100156. [CrossRef]
- Liu, X.; Qiu, T.; Dai, B.; Yang, L.; Liu, A.; Wang, J. Swarm-Intelligence-Based Rendezvous Selection via Edge Computing for Mobile Sensor Networks. *IEEE Internet Things J.* 2020, 7, 9471–9480. [CrossRef]
- 74. Cao, J.; Zhang, X.; Zhang, C.; Feng, J. Improved convolutional neural network combined with rough set theory for data aggregation algorithm. *J. Ambient Intell. Humaniz. Comput.* **2020**, *11*, 647–654. [CrossRef]
- Sanjay Gandhi, G.; Vikas, K.; Ratnam, V.; Suresh Babu, K.; Gandhi, G.S.; Vikas, K.; Ratnam, V.; Babu, K.S. Grid clustering and fuzzy reinforcement-learning based energy-efficient data aggregation scheme for distributed WSN. *IET Commun.* 2020, 14, 2840–2848. [CrossRef]
- Soundari, A.G.; Jyothi, V.L. Energy Efficient Machine Learning Technique for Smart Data Collection in Wireless Sensor Networks. *Circuits Syst. Signal Process.* 2020, 39, 1089–1122. [CrossRef]
- 77. Lyu, Z.; Wei, Z.; Wang, X.; Fan, Y.; Xia, C.; Shi, L. A Periodic Multinode Charging and Data Collection Scheme With Optimal Traveling Path in WRSNs. *IEEE Syst. J.* 2020, *14*, 3518–3529. [CrossRef]
- Wan, R.; Xiong, N.; Hu, Q.; Wang, H.; Shang, J. Similarity-aware data aggregation using fuzzy c-means approach for wireless sensor networks. *EURASIP J. Wirel. Commun. Netw.* 2019, 2019, 59. [CrossRef]
- Raja Basha, A.; Yaashuwanth, C. Optimal partial aggregation based energy delay compromise technique for wireless sensor network. *IETE J. Res.* 2019, 65, 855–871. [CrossRef]
- Mehrjoo, S.; Khunjush, F. Optimal data aggregation tree in wireless sensor networks based on improved river formation dynamics. Comput. Intell. 2018, 34, 802–820. [CrossRef]
- Zhang, J.; Hu, P.; Xie, F.; Long, J.; He, A. An energy efficient and reliable in-network data aggregation scheme for WSN. *IEEE Access* 2018, 6, 71857–71870. [CrossRef]
- Qadori, H.Q.; Zukarnain, Z.A.; Hanapi, Z.M.; Subramaniam, S. FuMAM: Fuzzy-based mobile agent migration approach for data gathering in wireless sensor networks. *IEEE Access* 2018, 6, 15643–15652. [CrossRef]
- Ghosh, N.; Banerjee, I.; Sherratt, R.S. On-demand fuzzy clustering and ant-colony optimisation based mobile data collection in wireless sensor network. *Wirel. Netw.* 2019, 25, 1829–1845. [CrossRef]
- Izadi, D.; Ghanavati, S.; Abawajy, J.; Herawan, T. An alternative data collection scheduling scheme in wireless sensor networks. Computing 2016, 98, 1287–1304. [CrossRef]
- Nithyakalyani, S.; Kumar, S.S. Voronoi fuzzy clustering approach for data processing in WSN. Int. J. Comput. Intell. Syst. 2014, 7, 105–113. [CrossRef]
- Zhu, Y.h.; Wu, W.d.; Pan, J.; Tang, Y.p. An energy-efficient data gathering algorithm to prolong lifetime of wireless sensor networks. *Comput. Commun.* 2010, 33, 639–647. [CrossRef]
- 87. Zhu, R.; Qin, Y.; Wang, J. Energy-aware distributed intelligent data gathering algorithm in wireless sensor networks. *Int. J. Distrib. Sens. Netw.* **2011**, *7*, 235724. [CrossRef]
- 88. Dhivya, M.; Sundarambal, M. Cuckoo search for data gathering in wireless sensor networks. *Int. J. Mob. Commun.* 2011, 9, 642–656. [CrossRef]
- Osamy, W.; Salim, A.; Khedr, A.M.; El-Sawy, A.A. IDCT: Intelligent Data Collection Technique for IoT-Enabled Heterogeneous Wireless Sensor Networks in Smart Environments. *IEEE Sens. J.* 2021, 21, 21099–21112. [CrossRef]

- Osamy, W.; Khedr, A.M.; El-Sawy, A.A.; Salim, A.; Vijayan, D. IPDCA: Intelligent Proficient Data Collection Approach for IoT-Enabled Wireless Sensor Networks in Smart Environments. *Electronics* 2021, 10, 997. [CrossRef]
- Bhushan, S.; Kumar, M.; Kumar, P.; Stephan, T.; Shankar, A.; Liu, P. FAJIT: A fuzzy-based data aggregation technique for energy efficiency in wireless sensor network. *Complex Intell. Syst.* 2021, 7, 997–1007. [CrossRef]
- Sajedi, S.N.; Maadani, M.; Moghadam, M.N. F-LEACH: A fuzzy-based data aggregation scheme for healthcare IoT systems. J. Supercomput. 2021, 78, 1030–1047. [CrossRef]
- 93. Saranraj, G.; Selvamani, K.; Malathi, P. A novel data aggregation using multi objective based male lion optimization algorithm (DA-MOMLOA) in wireless sensor network. *J. Ambient Intell. Humaniz. Comput.* **2021**, 1–9. [CrossRef]
- Sarode, P.; Reshmi, T.; Pattabiraman, V. Combination of Fitness-Mated Lion Algorithm with Neural Network for Optimal Query Ordering Data Aggregation Model in WSN. *Wirel. Pers. Commun.* 2021, 116, 513–538. [CrossRef]
- Ajmi, N.; Helali, A.; Lorenz, P.; Mghaieth, R. MWCSGA—Multi Weight Chicken Swarm Based Genetic Algorithm for Energy Efficient Clustered Wireless Sensor Network. *Sensors* 2021, 21, 791. [CrossRef] [PubMed]
- Yun, W.K.; Yoo, S.J. Q-Learning-Based Data-Aggregation-Aware Energy-Efficient Routing Protocol for Wireless Sensor Networks. IEEE Access 2021, 9, 10737–10750. [CrossRef]
- Maivizhi, R.; Yogesh, P. Q-learning based routing for in-network aggregation in wireless sensor networks. Wirel. Netw. 2021, 27, 2231–2250. [CrossRef]
- Wei, X.; Yan, S.; Wang, X.; Guizani, M.; Du, X. STAC: A spatio-temporal approximate method in data collection applications. *Pervasive Mob. Comput.* 2021, 73, 101371. [CrossRef]
- 99. Bajpai, P.; Kumar, M. Genetic algorithm—An approach to solve global optimization problems. *Indian J. Comput. Sci. Eng.* **2010**, *1*, 199–206.
- 100. El-Sawy, A.A.; Hussein, M.A.; Zaki, E.; Mousa, A. An introduction to genetic algorithms: A survey a practical issues. *Int. J. Sci. Eng. Res.* 2014, *5*, 252–262.
- Abdmouleh, Z.; Gastli, A.; Ben-Brahim, L.; Haouari, M.; Al-Emadi, N.A. Review of optimization techniques applied for the integration of distributed generation from renewable energy sources. *Renew. Energy* 2017, 113, 266–280. [CrossRef]
- Nguyen, T.T.; Nguyen, N.D.; Nahavandi, S. Deep reinforcement learning for multi-agent systems: A review of challenges, solutions and applications. arXiv 2018, arXiv:1812.11794.