A Systematic Review of Data Aggregation Techniques in Wireless Sensor Networks

Iman Dakhil Idan Saeedi¹,⁴, Ali Kadhum M. Al-Qurabat¹,⁴, *

¹Department of Computer Science, College of Science for Women, University of Babylon, Babylon, Iraq.

E-mail: ¹iman.idan@student.uobabylon.edu.iq, *Corresponding author: alik.m.alqurabat@uobabylon.edu.iq

Abstract. A Wireless Sensor Network (WSN) consists of a number of sensor nodes which can sense, communicate, and store data, and it has a battery capacity. Data aggregation can be defined as a procedure applied for the elimination of redundant transmissions, and it provides fused information to the base stations, that in turn improve the energy effectiveness and increases the life span of energy constrained WSNs. The current article presents a systematic review of data aggregating schemes in flat and hierarchical WSNs, which are compared in light of different factors, including the node heterogeneity, mobility of aggregator nodes, and types of networks and algorithms. By comparing these remarkable techniques, a set of aspects are suggested to be dealt with in future works.

Index words: Data aggregation, WSNs, network lifetime, Flat Network, Hierarchical network.

1. Introduction

Wireless sensor networks (WSNs) are found to be applied in a number of applications, such as military surveillance, and the monitoring of facilities and environments [1]. A WSN often consists of several sensor nodes among which communication is enabled, in addition to the fact that these could communicate to any external sinks or base stations [2]. These sensors are distributed either in a random manner over a harsh environment (like a battlefield), or are purposely located in certain places. The coordination among these sensors forms different communication networks, including single and multi-hop networks, or a hierarchically organized structure that has a number of clusters and cluster heads [3]. The data is periodically sensed, processed, and transmitted by the sensor towards the base stations. How often data is reported, as well as how many sensors are involved in the process, are both determined by the particular application [4]. Data gathering can be described as the process of systematically collecting data that was sensed by means of a number of sensors, so that it will be sent to the base station to be processed. Given the fact that the sensor node is constrained in terms of energy, the direct transmission of data to the base station by all
sensors would rather be of no efficiency [5]. This is due to the redundancy of data that resulted from neighbour sensors, which is of rather high correlation. Furthermore, base stations cannot process the huge amounts of data that are generated within a larger sensor network [6]. Therefore, certain networks are required so as to combine data and create significant information at sensors or intermediate nodes, as this could contribute to the reduction of packet transmissions to the base station and thus conserve energy and bandwidth. Such a case could be realized by means of applying data aggregating processes, as these usually involve the fusion of data obtained through more than one sensor at intermediate nodes and transmitting aggregated data to base stations (sink) [7].

This paper aims to contribute to the following aspects:

- Systematically reviewing and examining the current approaches of data aggregation in WSNs.
- Exploring initial obstacles and difficulties of data aggregating implementations within WSNs.
- Dividing data aggregation techniques into two categories, namely Flat Network and Hierarchical Network.
- Providing an accurate assessment for open issue and future trends.

The remainder of this article is organized in the following way: in Section 2, a general overview of data aggregation is presented. In Section 3, the various data aggregating protocols are classified according to the network structures that are related to the aggregation process. Section 4 summarized the comparison drawn between the related studies, whereas Section 5 states the concluding remarks and refers to possible future works. http://conferenceseries.iop.org/mse

2. Data Aggregation Overview

Data aggregation can be described as the procedure of data that is aggregated from more than one sensor, for the purpose of eliminating redundant transmissions and providing fused information to the base stations [6]. Fasolo et al. (2007) defined the process of data aggregating as, “the global process of gathering and routing information through a multi-hop network, processing data at intermediate nodes with the objective of reducing resource consumption (in particular energy), thereby increasing network lifetime.”

Figure 1 (A) shows the upward movement of data packets within the hierarchical structure, causing the overhead to increase. Transmitting redundant sensor readings imposes a huge communicating overhead upon the sensor nodes that are located near the base stations. In addition, the energy balance is lost between nodes at several levels within the hierarchical structure, mainly due to the converge-cast communication that takes place. The loss of energy balance in turn decreases the life span of the nodes, so that the ones closer to the base station will have a shorter duration than leaf nodes. Figure 1 (B), on the other hand, illustrates the drastic reduction of communicating traffic at intermediate nodes whenever en-routed data aggregation is applied [8].
The data aggregation process tries to obtain critical data through sensors, so as to provide it to the sink energy-efficiently, while keeping the data latency minimal. The latter is significant in a variety of uses, including environmental monitoring, which considers data freshness of high importance [9]. Developing energy-effective data aggregating algorithms is crucial for enhancing the life span of networks. The general data aggregation algorithm functions as shown in Figure 2.

A number of factors may influence how energy-efficient sensor networks are, including its structure, the data aggregating mechanisms, and the used routing protocols. The current study describes the effect of each factor on the network's energy efficiency in such a case.

- **Energy Efficiency**: This concept could be formally defined as the functionality of sensor networks which extend for the longest time possible. Within ideal data aggregating schemes, all sensors need to have expended the same amount of energy in each data gathering round. A data aggregation scheme is energy efficient if it maximizes the functionality of the network. Assuming that all sensors are equally important, the energy consumption of each sensor should be minimalized. This idea is captured by the network lifetime which quantifies the energy efficiency of the network [1,3].

There are a number of crucial factors to measure the performances of data aggregating algorithms, such as the life span of networks, data accuracy and latency. Taking into
consideration that these measures are defined differently according to their applications, these concepts can be formally defined as follows:

- **Network lifetime**: It is a certain number specified by system designers. To exemplify, whenever the application is determined by the time of node cooperation, the life time is described as the number of rounds performed before draining the energy of the 1st sensor [4].

- **Data accuracy**: This concept is defined according to the particular applications that the sensor networks are designed for [4].

- **Latency**: It is the delay that occurs in terms of transmitting data, routing, and aggregating data. Its measurements take form of the time delay that takes place between receiving data packets at the sink and generating data at source nodes [4].

In general, there are a number of benefits of applying data aggregating approaches, as well as several disadvantages and challenges in comparison without data aggregation as illustrated in Table 1.

**Table 1.** A comparison between the benefits, drawback, and difficulties of scenarios whereby data aggregation is performed versus without.

|                         | Without Aggregation                                                                 | With Aggregation                                                                 |
|-------------------------|-------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| **Advantages**          | It can be applied to all environmental measurements.                                | It increases how efficient and accurate the information is that has been obtained overall throughout the network. |
|                         | Routing issues could be easily solved using simple linear programs.                | A particular form of redundant data is found in the data collected via the nodes, so the role of this procedure is to eliminate any information that is not necessary. |
|                         |                                                                                     | A decrease in traffic load and energy consumption is observed.                  |
| **Disadvantages**       | It has an increased amount and size of data in transmissions.                      | It cannot be applied to all types of measuring environments.                   |
|                         | Transmitting data between nodes is rather energy consuming.                      | The aggregator or cluster head could witness attacks or failure.               |
|                         | More than one node could receive the same data.                                   | The amount of energy consumed might increase occasionally.                    |
|                         |                                                                                     | Collecting data increases the energy consumption and thus requires the removal of duplicate data to regain its energy efficiency |
| **Challenges**          | The probability of having a cavity formed close to the sink.                      | Data aggregation needs to be performed within limited time intervals.          |
|                         | Networks could outage over longer distance spans when transmitting data of high volume, as some data is taken by certain nodes. | Varying compression rates have to be carried out according to its application. |
|                         |                                                                                     | The energy consumption needs to be optimized, due to the limitation in battery power of nodes and timing protocols to switch into sleep-mode. |
|                         |                                                                                     | To increase the efficiency, unused sensor nodes could be made use of when in storing data. |
3. Data Aggregation Protocols According to Network Architecture

In this section, it will be explained how the data aggregation depends on the types of the network, which can be classified into two categories, namely flat network and hierarchical network. Each type will be elaborated on in detail.

3.1. Flat Networks

Flat networks have a crucial function within WSNs, as all sensor nodes have batteries that are equally power-supplied, and function similarly within the networks. This type of networks requires data centrically routing in data aggregation, whereas the sink transmits data packets to sensor nodes, as in floods. Flooding sensors carry data that matches the data packets and sends the response data packets back to the sink [10].

3.2. Hierarchical Networks

The high rates of energy consumption in flat networks is often due to the fact that any form of communicating and computing create a burden upon the sink. As for hierarchical networks, the data is aggregated by means of a special node which reduced the number of data packets sent to the sink. Therefore, such a structure improves the overall energy efficient performance of the hierarchical networks, which consist of Centralized, In-Network, Tree Based and Cluster Based [10]. Table 2 illustrated a number of significant differences between the flat and hierarchical data aggregating protocol.

| **Hierarchical network** | **Flat network** |
|--------------------------|-----------------|
| A cluster head or leader node aggregates the data. | Different nodes along the multi-hop path may aggregate the data. |
| The overhead influences the creation of clusters or chains within networks. | The routes of data aggregation routes can only be created within regions with transmittable data. |
| Networks remain operational whenever one of the cluster heads fail. | The whole network may break down whenever the sink node fails. |
| The short-range transmissions between sensor nodes and cluster heads results in a low latency rate. | The data transmitted to the sink along the multi-hop paths result in a high latency rate. |
| The routing structures are generally straightforward but not always optimal. | The additional overhead guarantees optimal routing structures. |
| The assignment of a higher-energy node as cluster head exploits the node heterogeneity. | No node heterogeneity is used in the improvement of energy effectiveness. |

3.2.1. Centralized Approach

As for this type of approaches, all sensors send their sensed data in form of data packets to a central node or base station through the briefest route available. The role of the aggregator or header node is to aggregate the data received via the other nodes, after which the resulting data is transmitted in form of a single packet. [11]. Direct Diffusion (DD) [12] and Sensor
Protocol for Information via Negotiation (SPIN) [13] are some examples of this type of protocol. Figure 3 illustrates the centralized aggregation approach.

3.2.2. In-Network Approach
In-network aggregating is a thorough approach to gather and process data at intermediate nodes, as well as to route information via multi-hop networks. It mainly aims to reduce the amount of power consumed throughout the process [11]. Two forms of in-network aggregation are known:

1. Size-reducing aggregation: The data packets coming from sensor nodes through their neighbour nodes are combined and compressed to reduce the size of packets to be sent to the sink node.
2. With no size reduction: The value of data is not processed here when merging the packets of various neighbour nodes into one packet.

Some examples of this type of protocol are Data Routing In-Network Aggregation (DRINA) [14] and Modified Data Routing In-Network Aggregation (M-DRINA) [15].

3.2.3. Tree Based Approach
Data Aggregation Trees (DAT) are created whereby all data transmissions require the construction of minimal spanning trees. All nodes within the network have a parent-child relationship where data is directed in a bottom-up approach, as shown in Figure 5. The data flows starting by the leaf nodes towards the sink node, while the parent nodes aggregate the data within the networks [11]. An example of this type of protocol is Tiny Aggregation (TAG) [16].
3.2.4. Cluster Based Approach
This approach involves splitting networks into a number of clusters. Every cluster consists of a set of sensor nodes, one of which is selected to be the cluster head. The cluster head performs the role of data aggregation whereby the data obtained gets aggregated and is then forwarded to the sink. The bandwidth here is brought to a minimum due to the reduced number of packets to be sent. In this approach, the data aggregating process serves to reduce any directly transmitted packets to the base station, in addition to the decrease in the amount of energy consumed as a result of the reduced distance of transmission [11]. There are a number of cluster-based approaches which have been introduced for wireless sensor networks, such as Low Energy Adaptive Clustering Hierarchy (LEACH) [17], Hybrid Energy Efficient Distributed Clustering Approach (HEED) [18] and clustered diffusion with dynamic data aggregation (CLUDDA) [19].

4. Summary and Comparison Works
After having discussed and compared the different works related to the topic in concern, this section summarizes the works referred to in the survey, as shown in Table 3. Four main key criteria have been taken into account when surveying the related works, namely the type of nodes, networks, and algorithms, as well as the mobility of aggregators.

1. Node types: There are two types of nodes found in wireless sensor networks: homogeneous and heterogeneous. The latter is characterized by more advanced abilities, and the one that has been taken into consideration here is its capability of increased initial energy.
2. Network types: Network topology is highly significant in exploiting redundant data. The networks can be either clustered, tree-based, or flat without node groupings.

3. Algorithms: There are three types of aggregating algorithms, namely central, distributed, and local algorithms. In the first type, data is coordinated and aggregated by a central node itself or it assigns another aggregator. As for the second, the distributed operation of the algorithm creates an equal likelihood for all nodes to take up the role of aggregator, based on a particular selecting criterion. Finally, with the third type the algorithm functions locally and decides who aggregated the data as well as the manner of aggregation.

4. Aggregator mobility: The majority of initial studies mostly centred around the aggregator having a fixed place, and assumed that any of the nodes could function as such. Later on, approaches were added whereby aggregators are mobile nodes and not resource-constrained, enabling it to move and collect data at different locations within the network.

Certain remarks were presented for each study included in the survey in Table 3. These notes shed light on the approaches following each respective work.

Table 3. A summary of data aggregating mechanisms and their main characteristics.

| Paper | Year | Network Type | Node Type | Aggregator | Algorithm | Application Type | Remarks |
|-------|------|--------------|-----------|------------|------------|------------------|---------|
| [1]   | 2020 | Cluster      | Homogeneous | Fixed      | Distributed| Environment monitoring | SAX symbolic algorithm and adaptive piecewise constant approximation (APCA) as a data aggregation technique to get rid of the redundant data readings. |
| [20]  | 2020 | Cluster      | Homogeneous | Fixed      | Distributed| Environment monitoring | A clustering-based Dynamic Time Warping (DTW) in Fog-computing and in the sensor nodes an integrated grouping and simple encoding algorithm is applied. |
| [21]  | 2020 | Cluster      | Homogeneous | Fixed      | Distributed| Disaster Management and rescue operation | A cluster-based systematic data aggregation model (CSDAM) for real-time data processing |
| [22]  | 2020 | Cluster      | Homogeneous/heterogeneous | Fixed | Distributed | Environment monitoring | Provide a new data aggregation method based on the open-pit mining idea efficiently. |
| [23]  | 2020 | Cluster      | Homogeneous | Fixed | In-Network | Environment monitoring | Groups the homogeneous data into clusters and then performs in-network data reduction by selecting the central values of each cluster. |
| [24]  | 2019 | Cluster      | Homogeneous | Fixed | In-Network | Environment monitoring | Local aggregation algorithm, make use of the temporally correlating features in wireless Sensor Networks in eliminating data redundancy and local outlier data, so as to improve data quality and transmission ratios. |
| Ref. | Year | Cluster | Homogeneity | Network Type | Environment Monitoring | Description |
|------|------|---------|-------------|--------------|------------------------|-------------|
| [3]  | 2019 | Cluster | Homogeneous | Fixed, Distributed | Environment monitoring | Use Integrated Divide and Conquer with enhanced K-means technique for energy saving data aggregation in WSNs. |
| [25] | 2019 | Tree | Homogeneous | Fixed, Distributed | Wireless water meter application | Multi-channel TDMA scheduling algorithms are used in reducing collisions, in addition to the meta-heuristic algorithm to minimize energy and reduce latency. |
| [26] | 2019 | Cluster | Homogeneous | Mobile, Distributed | Disaster-prone zone monitoring | Neural network-based cluster formation and Ant Colony based optimization |
| [27] | 2019 | Flat | Homogeneous | Fixed, Distributed | Monitoring applications | The aggregator makes use of a linear classifier on a basis of SVM To identify and eliminate data redundancy. |
| [28] | 2018 | Flat | Homogeneous | Mobile, Distributed | Monitoring applications | Adaptive Particle Swarm Optimization is used for optimum path selection of the mobile node |
| [7]  | 2018 | Cluster | Homogeneous | Fixed, Distributed | Environment monitoring | Data dimensionality is aggregated and used through an Adaptive Piece-wise Constant Approximating method. |
| [29] | 2018 | Cluster | Homogeneous | Fixed, Distributed | Precision Agriculture applications | During consecutive iterations transmission of redundant data is reduced by using differential data from sensor. |
| [6]  | 2018 | Cluster | Homogeneous | Fixed, Distributed | Environment monitoring | Remove the redundant collected data and adapts its sampling rate in accordance with the monitored environment conditions. |
| [30] | 2017 | Cluster | Homogeneous | Fixed, Distributed | Environment monitoring | An efficient two phases of data aggregation, in the first phase, similarities between measures collected by each sensor used. In the second phase, it uses distance-based function to find similarity between sets of collected data. |
| [31] | 2017 | Cluster | Homogeneous | Fixed, Distributed | Environment monitoring | Sensor readings are collected periodically, and the lifespan is extended in Periodic Sensor Networks (PSN). |
| [32] | 2016 | Cluster | Homogeneous | Fixed, Distributed | Terrain environment monitoring | The formation of the clustered GSTEB tree takes place within the cluster, and the root of the tree is the CH. It has a better throughput than GSTEB. |
| [33] | 2016 | Flat | Homogeneous | Fixed, Centralized | Environment monitoring | Two levels: the first filters the data through the fitting function, when the norm value of data is below a threshold, then nothing is forwarded to the aggregator to be aggregated. |
| [34] | 2016 | Cluster | Homogeneous | Fixed, Distributed | Environment monitoring | Use pre-filtration techniques on nodes data pre-processing can |
suppress a huge amount of redundant or correlated data transmissions, thus maximizes the node battery lifetime.

A three-layer big data aggregation framework using priority-based dynamic data aggregation scheme (PDDA)

Trade-offs between data quality and energy dissipation for realizing an efficient data aggregation. There are two algorithm types: Distributed Energy Allocation (DBA) and Centralized Energy Allocation. The GA based approach and Gibbs Sampling Theory are used.

It has mobile link and rendezvous node which functions as a storage for mobile link.

It uses data density correlation, represented by the spatial correlation between node data and its neighbors.

The framework is based on energy replenishment and anchor-based data gathering with mobile aggregators.

Energy-balanced Transmission Protocol (ETP) is proposed which uses the slice-based energy model. The balance is maintained through inter- and intra-slice energy balancing.

The coordinator node (CN) is chosen in a random manner in a cluster. The CH transmits data to CN, which performs loss ratio calculation and modifies the clusters if necessary.

Reducing energy consumption.

Data is collected by the mobile nodes and uses meta-heuristic approach for selecting polling points.

Does an in-network aggregation of data based on the order compression techniques. Based on history, data may be suppressed.

Multiple nodes, which are mobile used Multi mobile sink based heuristic approach for division of

| Year | Type       | Network        | Data Quality | Application                                                                 |
|------|------------|----------------|--------------|-----------------------------------------------------------------------------|
| 2016 | Cluster,   | Homogeneous    | Distributed  | Emergency/Critical operation
|      | Tree       |                |              | A three-layer big data aggregation framework using priority-based dynamic data aggregation scheme (PDDA) |
| 2015 | Cluster,   | Heterogeneous  | Centralized, | Monitoring Applications
|      | Flat       |                | Local        | Trade-offs between data quality and energy dissipation for realizing an efficient data aggregation. There are two algorithm types: Distributed Energy Allocation (DBA) and Centralized Energy Allocation. The GA based approach and Gibbs Sampling Theory are used. |
| 2015 | Cluster    | Homogeneous    | Mobile       | Distributed Health and underwater monitoring It has mobile link and rendezvous node which functions as a storage for mobile link. |
| 2014 | Tree       | Homogeneous    | Fixed        | Distributed Monitoring Applications It uses data density correlation, represented by the spatial correlation between node data and its neighbors. |
| 2013 | Tree       | Homogeneous    | Mobile       | Distributed Monitoring Applications The framework is based on energy replenishment and anchor-based data gathering with mobile aggregators. |
| 2012 | Tree       | Homogeneous    | Fixed        | Distributed Environment monitoring Energy-balanced Transmission Protocol (ETP) is proposed which uses the slice-based energy model. The balance is maintained through inter- and intra-slice energy balancing. |
| 2012 | Cluster    | Homogeneous    | Distributed, | Environment monitoring The coordinator node (CN) is chosen in a random manner in a cluster. The CH transmits data to CN, which performs loss ratio calculation and modifies the clusters if necessary. |
| 2012 | Cluster    | Homogeneous    | Distributed, | Environment monitoring |
| 2012 | Cluster    | Homogeneous    | Distributed  | Health monitoring Reducing energy consumption. |
| 2011 | Tree       | Homogeneous    | Mobile       | Distributed Mobile Target tracking Data is collected by the mobile nodes and uses meta-heuristic approach for selecting polling points. |
| 2011 | Tree       | Homogeneous    | Fixed        | In-Network Environment monitoring Does an in-network aggregation of data based on the order compression techniques. Based on history, data may be suppressed. |
| 2010 | Tree       | Homogeneous    | Mobile       | Distributed Mobile Target tracking Multiple nodes, which are mobile used Multi mobile sink based heuristic approach for division of |
| Year | Type | Model | Level | Environment | Data Management | Description |
|------|------|-------|-------|-------------|-----------------|-------------|
| 2009 | Cluster | Homogenous | Fixed | Distributed | Health monitoring | In V-LEACH protocol, besides having a CH in the cluster, there is a vice-CH that takes the role of the CH when the CH dies. |
| 2006 | Tree | Homogenous | Fixed | Distributed | Health monitoring | A different approach for collaboration among the nodes is used and construct a tree-like hierarchical path of the nodes through which the data is aggregated. |
| 2004 | Cluster | Heterogeneous | Fixed | Distributed | Environment monitoring | The approach is hybrid: cluster heads are probabilistically selected based on their residual energy, and nodes join clusters such that communication cost is minimized. Cluster heads are responsible for aggregation of their data. |
| 2003 | Cluster | Heterogeneous | Mobile | Centralized | Health monitoring | Combined Directed Diffusion with clustering to prevent flooding during interest propagation and Dynamic data aggregation points. |
| 2002 | Tree | Homogenous | Fixed | Distributed | Environment monitoring | TAG consists of two phases: a distribution phase, in which aggregate queries are pushed down into the network, and a collection phase, where the aggregate values are continually routed up from children to parents. |
| 2002 | Chain | Homogenous | Fixed | Distributed | Monitoring | The main idea is for each node to receive from and transmit to close neighbors and take turns being the leader for transmission to the BS. |

5. Open Issue and Future Trends

There are a number of significant issues and obstacles that need to be addressed in future works. This section discusses and investigates the main ones among them. None of the techniques could cover all of the problems faced in data aggregation in IoT, as some take into account factors like the life span of networks and the amount of energy consumed, whereas others tend to neglect these aspects. In addition, the evaluation of some methods was done by means of simulations, whereas others made use of real-world scenarios for obtaining realistic results.

A remarkable challenge is improving mechanisms with more than one sink, and it might require further research. The combination of mechanisms and energy-aware routing or MAC layer protocols for extending the lifespan of IoT networks is also an important aspect to be examined. Other aspects to explore are the reduction of overhead computing costs, ensuring a better security against attacks, and the use of suitable protocols within cybersecurity, such as risks, threats, and managing vulnerable aspects. Such issues need to be taken into account before turning to practically constructed architectures.
The cluster-based approach is highly efficient within static networks, as no changes occur to their structures over a longer period of time. Yet, such a technique tends to be of no efficiency when applied within a dynamic network. It would therefore be useful to examine the ways of data aggregating for dynamic networks. Another concept to be dealt with is the impact of sensor node heterogeneity on the aggregating techniques. In addition, techniques whereby stronger sensor nodes aggregate the data indicate promising results, yet the determination of their position might need further researching.

On a closing note, an unquestionably important aspect in data aggregation is its security. There are indeed studies that have dealt with this aspect but have not been addressed in this survey. This is due to the fact that such issues are accompanied with the repeated incorrect transmitting of data by aggregators. It is particularly this notion that has not been addressed enough in research, as all available techniques depend on extensive node monitoring mechanisms. The extent to which such node monitoring techniques are efficient has not been examined sufficiently, as they are affected by high radios and consumed sensing resources. As such, the expanded light-weight monitoring mechanisms, particularly when providing security during data aggregating processes, represents a remarkable phenomenon to be studied in future works.

6. Conclusion
Wireless sensor networks are networks that have an energy constraint. Given that the majority of energy is used in the transmission of data, it is necessary to apply a form of data aggregating. This process eliminates duplicate data transmissions within the network, and as a result it has gained a considerable amount of interest recently. This paper presented a survey of various methods in data aggregating, based on a number of factors, including the types of networks and algorithms, node heterogeneity, and the aggregator mobility. The findings indicate that all techniques attempted to enhance the energy efficiency and latency. Tree-based data aggregating seems to be more favourable than the clustered approach, due to its reduction in latency and amount of energy consumed, but suffers from a high overhead when creating tree topologies. Cluster-based mechanisms reduced the amount of energy dissipated and expanded the scalability of networks. The centralized networks were least preferable due to their overhead and limited network dynamicity.

7. References

[1] Al-Qurabat, A.K.M. and Kadhum Idrees, A., 2020, Data gathering and aggregation with selective transmission technique to optimize the lifetime of Internet of Things networks. International Journal of Communication Systems, 33(11), p.e4408.

[2] Abdulzahra, S.A., Al-Qurabat, A.K.M. and Idrees, A.K., 2020, March. Data Reduction Based on Compression Technique for Big Data in IoT. In 2020 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 103-108). IEEE.

[3] Idrees, A.K., Al-Qurabat, A.K.M., Abou Jaoude, C. and Al-Yaseen, W.L., 2019, June. Integrated divide and conquer with enhanced k-means technique for energy-saving data aggregation in wireless sensor networks. In 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC) (pp. 973-978). IEEE.

[4] Al-Qurabat, A.K.M. and Idrees, A.K., 2019. Two level data aggregation protocol for prolonging lifetime of periodic sensor networks. Wireless Networks, 25(6), pp.3623-3641.
[5] Idrees, A.K. and Al-Qurabat, A.K.M., 2017. Distributed Adaptive Data Collection Protocol for Improving Lifetime in Periodic Sensor Networks. IAENG International Journal of Computer Science, 44(3).

[6] Al-Qurabat, A.K.M. and Idrees, A.K., 2018. Energy-efficient adaptive distributed data collection method for periodic sensor networks. International Journal of Internet Technology and Secured Transactions, 8(3), pp.297-335.

[7] Al-Qurabat, A.K.M. and Idrees, A.K., 2018. Distributed data aggregation and selective forwarding protocol for improving lifetime of wireless sensor networks. Journal of Engineering and Applied Sciences, 13(5 S1), pp.4644-4653.

[8] Fasolo, E., Rossi, M., Widmer, J. and Zorzi, M., 2007. In-network aggregation techniques for wireless sensor networks: a survey. IEEE Wireless Communications, 14(2), pp.70-87.

[9] Sangolgi, N.B. and Zakir, S.K.A., 2013. Energy aware data aggregation technique in WSN. International Journal of Scientific and Research Publications, 3(10).

[10] Mishra, S. and Thakkar, H., 2012. Features of WSN and Data Aggregation techniques in WSN: A Survey. Int. J. Eng. Innov. Technol.(IJETT), 1(4), pp.264-273.

[11] Shilpa, S.G. and Sundaram, S.M., 2017. Data Aggregation Techniques Over Wireless Sensor Network-A Review. International Journal of Engineering Research & Technology, 5(22).

[12] Intanagonwiwat, C., Govindan, R., Estrin, D., Heidemann, J. and Silva, F., 2003. Directed diffusion for wireless sensor networking. IEEE/ACM transactions on networking, 11(1), pp.2-16.

[13] Kulik, J., Heinzelman, W. and Balakrishnan, H., 2002. Negotiation-based protocols for disseminating information in wireless sensor networks. Wireless networks, 8(2-3), pp.169-185.

[14] Villas, L.A., Boukerche, A., Ramos, H.S., De Oliveira, H.A.F., de Araujo, R.B. and Loureiro, A.A.F., 2012. DRINA: A lightweight and reliable routing approach for in-network aggregation in wireless sensor networks. IEEE Transactions on Computers, 62(4), pp.676-689.

[15] Ujawe, P.V. and Khiani, S., 2014. Review on data aggregation techniques for energy efficiency in wireless sensor networks. International Journal of Emerging Technology and Advanced Engineering, 4(7).

[16] Madden, S., Franklin, M.J., Hellerstein, J.M. and Hong, W., 2002. TAG: A Tiny aggregation service for ad-hoc sensor networks. ACM SIGOPS Operating Systems Review, 36(5), pp.131-146.

[17] Heinzelman, W.B., Chandrakasan, A.P. and Balakrishnan, H., 2002. An application-specific protocol architecture for wireless microsensor networks. IEEE Transactions on wireless communications, 1(4), pp.660-670.

[18] Younis, O. and Fahmy, S., 2004. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Transactions on mobile computing, 3(4), pp.366-379.

[19] Chatterjea, S. and Havinga, P., 2003, October. CLUDDA-Clustered diffusion with dynamic data aggregation. In Ajaccio, Corsica, France: Cabernet Radicals Workshop (pp. 1-6).
[20] Idrees, A.K. and Al-Qurabat, A.K.M., 2020. Energy-Efficient Data Transmission and Aggregation Protocol in Periodic Sensor Networks Based Fog Computing. Journal of Network and Systems Management, 29(1), pp.1-24.

[21] Shobana, M., Sabitha, R. and Karthik, S., 2020. Cluster-based systematic data aggregation model (CSDAM) for real-time data processing in large-scale WSN. Wireless Personal Communications, pp.1-19.

[22] Ramezanifar, H., Ghazvini, M. and Shojaei, M., 2020. A new data aggregation approach for WSNs based on open pits mining. Wireless Networks, pp.1-13.

[23] Alam, M.K., Aziz, A.A., Latif, S.A. and Awang, A., 2020. Error-Aware Data Clustering for In-Network Data Reduction in Wireless Sensor Networks. Sensors, 20(4), p.1011.

[24] Verma, N. and Singh, D., 2019. Local Aggregation Scheme for Data Collection in Periodic Sensor Network. International Journal of Engineering and Advanced Technology, 9(22).

[25] Kumar, S. and Kim, H., 2019. Energy efficient scheduling in wireless sensor networks for periodic data gathering. IEEE access, 7, pp.11410-11426.

[26] Sarangi, K. and Bhattacharya, I., 2019. A study on data aggregation techniques in wireless sensor network in static and dynamic scenarios. Innovations in systems and software engineering, 15(1), pp.3-16.

[27] Yadav, S. and Yadav, R.S., 2019. Redundancy elimination during data aggregation in wireless sensor networks for IoT systems. In Recent Trends in Communication, Computing, and Electronics (pp. 195-205). Springer, Singapore.

[28] SreeRanjani, N.Y., Ananth, A.G. and Reddy, L.S., 2018. An energy efficient data gathering scheme in wireless sensor networks using adaptive optimization algorithm. Journal of Computational and Theoretical Nanoscience, 15(11-12), pp.3456-3461.

[29] Khriji, S., Raventos, G.V., Kammoun, I. and Kanoun, O., 2018, March. Redundancy elimination for data aggregation in wireless sensor networks. In 2018 15th International Multi-Conference on Systems, Signals & Devices (SSD) (pp. 28-33). IEEE.

[30] Harb, H., Makhouli, A., Layimani, D. and Jaber, A., 2017. A distance-based data aggregation technique for periodic sensor networks. ACM Transactions on Sensor Networks (TOSN), 13(4), pp.1-40.

[31] Al-Qurabat, A.K. and Idrees, A.K., 2017. Adaptive data collection protocol for extending lifetime of periodic sensor networks. Qalaai Zanist Journal, 2(2), pp.93-103.

[32] Shivkumar, S., Kavitha, A., Swaminathan, J. and Navaneethakrishnan, R., 2016. General self-organizing tree-based energy balance routing protocol with clustering for wireless sensor network. Asian Journal of Information Technology, 15(24), pp.5067-5074.

[33] Atouii, I., Ahmad, A., Medlej, M., Makhouli, A., Tawbe, S. and Hijazi, A., 2016, April. Tree-based data aggregation approach in wireless sensor network using fitting functions. In 2016 Sixth international conference on digital information processing and communications (ICDIPC) (pp. 146-150). IEEE.

[34] Kandukuri, S., 2016. Spatio-Temporal Adaptive Sampling Techniques for Energy Conservation in Wireless Sensor Networks (Doctoral dissertation).
[35] Miah, M.S. and Koo, I., 2012. Performance analysis of ILEACH and LEACH protocols for wireless sensor networks. Journal of information and communication convergence engineering, 10(4), pp.384-389.

[36] Xiao, S., Li, B. and Yuan, X., 2015. Maximizing precision for energy-efficient data aggregation in wireless sensor networks with lossy links. Ad Hoc Networks, 26, pp.103-113.

[37] Karim, L. and Al-kahtani, M.S., 2016, October. Sensor data aggregation in a multi-layer big data framework. In 2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (pp. 1-7). IEEE.

[38] Satapathy, S.S. and Sarma, N., 2006, April. TREEPSI: tree based energy efficient protocol for sensor information. In 2006 IFIP international conference on wireless and optical communications networks (pp. 4-pp). IEEE.

[39] Yassein, M.B., Khamayseh, Y. and Mardini, W., 2009, June. Improvement on LEACH protocol of wireless sensor network (VLEACH). In Int. J. Digit. Content Technol. Appl. 2009.

[40] Lindsey, S. and Raghavendra, C.S., 2002, March. PEGASIS: Power-efficient gathering in sensor information systems. In Proceedings, IEEE aerospace conference (Vol. 3, pp. 3-3). IEEE.

[41] Mottaghi, S. and Zahabi, M.R., 2015. Optimizing LEACH clustering algorithm with mobile sink and rendezvous nodes. AEU-International Journal of Electronics and Communications, 69(2), pp.507-514.

[42] Han, Z., Wu, J., Zhang, J., Liu, L. and Tian, K., 2014. A general self-organized tree-based energy-balance routing protocol for wireless sensor network. IEEE Transactions on Nuclear Science, 61(2), pp.732-740.

[43] Yuan, F., Zhan, Y. and Wang, Y., 2013. Data density correlation degree clustering method for data aggregation in WSN. IEEE Sensors Journal, 14(4), pp.1089-1098.

[44] Guo, S., Wang, C. and Yang, Y., 2013, April. Mobile data gathering with wireless energy replenishment in rechargeable sensor networks. In 2013 Proceedings IEEE INFOCOM (pp. 1932-1940). IEEE.

[45] Jin, N., Chen, K. and Gu, T., 2012, October. Energy balanced data collection in wireless sensor networks. In 2012 20th IEEE International Conference on Network Protocols (ICNP) (pp. 1-10). IEEE.

[46] Mathapati, B.S., Patil, S.R. and Mytri, V.D., 2012, September. Energy efficient reliable data aggregation technique for wireless sensor networks. In 2012 International conference on computing sciences (pp. 153-158). IEEE.

[47] Zhao, M., Ma, M. and Yang, Y., 2011. Efficient data gathering with mobile collectors and space-division multiple access technique in wireless sensor networks. IEEE Transactions on computers, 60(3), pp.400-417.

[48] Yang, C., Yang, Z., Ren, K. and Liu, C., 2011, October. Transmission reduction based on order compression of compound aggregate data over wireless sensor networks. In 2011 6th International Conference on Pervasive Computing and Applications (pp. 335-342). IEEE.

[49] Zhao, M. and Yang, Y., 2010, April. Data gathering in wireless sensor networks with multiple mobile collectors and SDMA technique sensor networks. In 2010 IEEE Wireless Communication and Networking Conference (pp. 1-6). IEEE.