Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 220 (2023) 187–194

Procedia Computer Science

www.elsevier.com/locate/procedia

The 14th International Conference on Ambient Systems, Networks and Technologies (ANT)
March 15-17, 2023, Leuven, Belgium

Designing A Secure and Long-Lived WSN for Data Collection

Wael Y. Alghamdi*[a]

[a]*Department of Computer Science , College of Computers and Information Technology, Taif University, P. O. Box 11099, Taif 21944, Saudi Arabia*

Abstract

Energy minimization is one of the significant tasks in the WSN which can be achieved through clustering process. Clustering in the WSN is an efficient method to increase the lifespan of the network. In the proposed work the main concentration is given to the network lifetime which is very necessary for the less node failures and high packet deliveries. In this paper the hierarchical structuring is done using clustering approach based on efficient data aggregation process. The aggregation process is defined into two phases. The first phase defines the arrangement of the nodes in the cluster using uniform distribution and second phase defines the selection of the cluster head and aggregation of the data using linear programming. The main objective is to minimize the energy consumption and increase of the robustness of the sensor network. Eventually the performance of the network will be evaluated in terms of low energy consumption, high packet deliveries and low congestions in the WSN.

Keywords: Network lifetime; clustering; wireless sensor network; linear programming;

1. Introduction

The Wireless Sensor Networks (WSNs) is a collection of a large number of tiny sensor nodes that sense the desired phenomena in respect of a particular region of interest and send the sensed data to a base stations. Sensor technology has a large number of applications, for example, disaster management, habitat monitoring, security and military intelligence building [1]. Sensor nodes are typically battery-powered, and the communication between sensor nodes and the base station consumes most of the energy. Accordingly, it is unrealistic to replace them once they run out of battery. Therefore, energy efficiency is essential for WSNs, and improving this is an important goal in maximising the lifetime of WSNs [2].

* Corresponding author. Tel.: +966 569441075
E-mail address: W.alghamdi@tu.edu.sa

10.1016/j.procs.2023.03.026

Clustering is a method used to increase the network lifetime and the robustness of WSNs effectively. In clustering, a WSN is partitioned into disjoint clusters, with each cluster having a cluster head (aggregator). Clustering decreases the overall energy consumption of the sensor network because the data size and communication sessions are reduced [3]. Data aggregation is a process of summarizing the data collected, effectively reducing the size of the data to be transmitted. The data aggregation happens when an aggregator collects the information from its neighbouring sensor nodes and sends it to the base station via a routing topology. Since the data size is significantly reduced after data aggregation, the energy consumed by transmitting and receiving data via sensor nodes is also significantly reduced [4].

In this paper, I investigate the problem of constructing clusters and routing data aggregation for the WSNs where each aggregator has two parents and I propose a novel approach. The proposed approach aims to minimize the maximize energy consumption. The proposed approach is divided into two phases: Clustering and routing construction heuristics and shortest path graph-based routing. The Approach has the following salient features:

- Using Homomorphic Encryption to ensure that no other sensor nodes can understand the content of data delivered to the base station and eliminate the costs for decryptions.
- Using clustering to improve the robustness of network topology and the network lifetime. Data aggregation is performed by each cluster head to reduce the data size and thus reduce the energy consumption for sending and receiving data.
- Using dual cluster heads to handle selective forwarding attacks and modification attacks. Each cluster head receives the aggregated data of all the descendant sensor nodes, and the two cluster heads of each cluster receive the same aggregated data.
- A tree-based routing topology. I propose an ILP(Integer Linear Programming) based approach to construct a routing tree aimed at maximising the network lifetime.
- A shortest path graph routing topology. I propose an LP (Linear Programming) based approach to construct a shortest path graph for routing such that the network lifetime is maximised.

The objective of this paper is to sustain security and energy efficiency within the wireless network design and at the same time maximize the network lifetime. However, this research presents some key challenges. Firstly, in order to handle attacks, I have to make use of data redundancy so that the base station receives redundant data in order to determine if the data received is correct. However, data redundancy decreases the network lifetime and complicates the network lifetime-aware routing algorithm. Secondly, partitioning all the sensor nodes into disjoint clusters so that all the clusters consume the same amount of energy is an NP-hard problem. Finally, the major goal of routing topology construction is to construct a routing topology for data delivery such that the network lifetime is maximized. This problem is NP-hard problem, and it is widely believed that no polynomial-time algorithm exists. As far as I know, previous research has not achieved both secure data aggregation and maximization of network lifetime.

2. Related Works

2.1. Clustering-Based Network Lifetime

Many network lifetime-aware clustering approaches have been proposed. Low Energy Adaptive Clustering Hierarchy (LEACH) is one of the leading clustering approaches in WSNs [5]. In LEACH, the network partitions the nodes into clusters and the cluster head is selected based on a predetermined probability threshold. Other nodes join the nearest cluster head to form clusters, which reduces the direct communication with the base station. In the data transmission, each cluster head aggregates data and sends the data to the base station directly. In this protocol, fast delivery is guaranteed, but it results in more energy being consumed for data transmission. [6] propose new clustering methods, such as distance, density and energy based clustering (DDEC). DDEC is used to improve the network performance, and it is classified into residual energy, distance and density. It partitions the network into cluster with the same member number to achieve the load balancing. The author finds that DDEC outperforms DDCHS in terms of alive node and energy consumption. [7] examines Data clustering in a wireless sensor network implemented on a self-organization feature map (SOFM) neura network. WSN is one of the promising communication networks for evaluating and monitoring the environmental area. In this type of network, communication is carried via radio signals. Hence, the sensor node has the capacity for data storage, sensing and processing.

Different studies have been conducted relating to network lifetime based on data aggregation. The study by [8] investigated adaptive aggregation routing with the goal of reducing the delay for multi-layer wireless sensor networks. The study found data aggregation to be a method that is very effective, especially in the reduction of the volume of data transmitted and improvement in network lifetime. Specifically, this study showed that when the wait time length is greater than or equal to the aggregation being predetermined, data is forwarded further but that this comes at the expense of delay. The AAR schemes in the study by [8] were therefore aimed at sending data to the nodes with long data queues with the purpose of ensuring that the long queues have a small amount of data so that the aggregation can be performed, thus reducing the amount of delay. All the above-mentioned aggregation protocols help in achieving those objectives as they work on the concept of successful data transmission with effective utilization of available resource parameters. Past studies have indicated that collaboration, coordination and management of the sensor nodes are necessary to eliminate the problem of energy inefficiency, lessening the network lifetime [9].

2.2. Network Lifetime Based Data Aggregation

Different studies have been conducted relating to network lifetime based data aggregation. One of these studies was carried out by [10] on mechanisms of data aggregation in IoT in which a systematic literature review was provided and future researchers were given recommendations. The study found that data aggregation is one of the most efficient methods used to decrease the transmission numbers among objects. The study also found that the loss in redundancy of data may contribute to the lengthening of the lifetime of the network as well as a decrease in energy consumption. [10] study provided systematic analysis of IoT in data aggregation mechanisms and found that there are three main mechanisms: clusterbased, tree-based and centralised. Also, a detailed comparison of significant techniques showed that the mechanism of data aggregation used in the IoT was critical to efficiency and loss of redundancy. With improvement in multimedia presentations, Sensor networks are now standard because of their valuable characteristics and varieties of the applications. Still, sensor networks contain very minor nodes and are dependent on the batteries and also the sequences of the nodes are not expendable. So the efficient work is done using high data aggregation control procedures using ant colony optimization and PSO to apply soft computing for the high network lifetime [11]. The nodes in the sensor networks are applied to gather the information using wireless medium. Information generated using neighboring nodes is extremely redundant causing high load of the traffic, delay, and other congestions which consumes high energy and reducing the network lifespan. Various data sources and an energy efficient method are worked on with low latency in terms of data transfer and reception [12]. As IOT is having huge demand. So data gathering is an important procedure to decrease the dissipation of the energy of network nodes to have high lifetime of the network. So data clustering and recognition of anomalous actions, a data aggregation using various clustering approaches like fuzzy cmeans is also in the market to reduce the redundancies for wireless sensor networks [13].

3. The Proposed Approach

3.1. Network Model

The target wireless sensor network consist of set n sensor nodes deployed on a 2D plane and the location of each sensor is unknown. All sensor nodes have the same initial energy and transmission range. I assume there is no communication barrier between any pair of sensors. Accordingly, every sensor node vi can communicate directly to another sensor node vj. The network is partitioned into k disjoint clusters with two aggregators for each cluster. Every node in the cluster having equal probability to become a CH and all elected cluster heads will act as relay nodes.

3.2. Clustering

1. Partition all the network into layers:

    - Draw circles $C_i (i = 1, 2, k)$ with an optimal radius. All the sensors between C_i and C_{i+1} belong to the same layer.

- Remove each layer without any sensor node and rename all the layers as $L_1, L_2, \ldots, L_m$ such that $C_i < C_{i+1} (i = 1, 2, \ldots, m-1)$.

2. Partition each layer into disjoint clusters

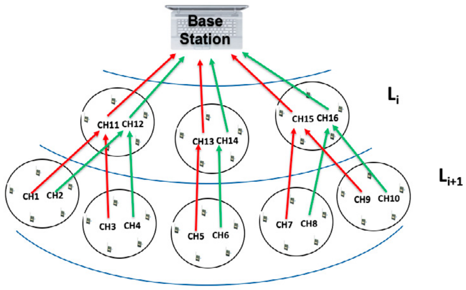

Fig. 1. An example for illustrating the algorithm Clusters

3.3. ILP Routing Tree Construction Heuristic

The objective of the ILP algorithm is to construct a shortest path spanning tree rooted at the base station such that the network lifetime is maximized. For each two adjacent layers L_{k+1} and $L_k (k = m-1, \ldots, 2, 1)$, do the following:

For each cluster v_j in L_{k+1} find the parent cluster v_i in L_k by solving an ILP below. Perform this step for each cluster to achieve minimum energy consumption for the high network lifetime. Perform the optimize Routing from v_j to v_i to achieve high network lifetime with low packet losses.

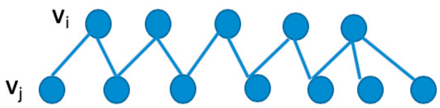

Fig. 2. An Instance for Two Layers

The following steps show the ILP for Two Layers L_k and L_{k+1} which consist of Binary variables constrains, Parent constrains, Data size constrains, energy constrains, and finally the objective function.

1. Binary variable: In order to represent each edge uniquely, I stipulate that an edge $E_{i,j}$ denotes the edge between cluster c_i in group G_{s+1} and cluster c_j in group G_s. For each edge $E_{i,j}$, I introduce a binary decision variable X_i;j as follows:
   For each edge (v_i, v_j) (v_i is in L_k and v_j is in (L_{k+1}), introduce a binary variable $x_{i,j}$

$$x_{i,j} = 1 \text{ if } v_i \text{ is the parent of } v_j \quad (1)$$

$$x_{i,j} = 0, \text{otherwise} \quad (2)$$

2. Parent constraints:

For every cluster v_j in L_{k+1}:

$$\sum_{v_i \in P(v_j)} x_{i,j} = 1 \tag{3}$$

The above constraint implies that among all the candidate parents of c_i, only one cluster can be selected as the parent where $P(v_j)$ is a set of possible parents of v_j.

3. Data size constraints:

$$W_i = F(w_i + \sum v_{ij} \in V_i \ x_{ij} * F(w_j)) \tag{4}$$

where W_i is the data size after aggregation in cluster v_i, w_i is the data size of cluster v_i , $\sum v_{ij} \in V_i \ x_{ij} * F(w_j)$ is the total data size of the data received from all children, $F(x)$ is the function mapping from the input data size to the output data size after aggregation. I consider two types:

(a) $F(x) = x$
(b) $F(x) = c$, where c is a constant

4. Energy constraints

$$E_i = e_i + \alpha \sum_{v_{ij} \in V} X_{ij} * F(w_j) + \beta * W_i + e(W_i) \tag{5}$$

where e_i (constant) is the energy spent in receiving the data from all members of cluster v_i, α is the energy spent in receiving one unit of data, β is the energy spent in transmitting one unit of data, and $e(W_i)$ is the energy in data aggregation by v_i. I can assume that $e(x)$ is a linear function $e(x) = k_x$, where k is a constant.

5. The optimization objective function is shown as follows:

$$min\ max\ E_i \tag{6}$$

3.4. Shortest Path Graph-based Routing

I use undirected connectivity graph to represent the wireless connectivity between sensor nodes in the WSN.

An undirected graph $G = (V, E, W)$, where V is a set of clusters, E is a set of edges where (v_i, v_j) denotes that cluster each sensor node in vi can communicate with each sensor node in v_j and v_i and v_j do not belong to the same layer, and each weight w_i in W denotes the size of cluster v_i.

After constructing the shortest path routing graph, I compute the flow on each edge of the graph. The flow on each edge represents the size of the aggregated data received from the edge. All the flows are computed such that the network lifetime is maximised. Next, I propose LP for calculating the optimal fows.

1. Data size constraints:

$$W_i = F(w_i + \sum_{v_j \in C(v_i)} f_{ij}) \tag{7}$$

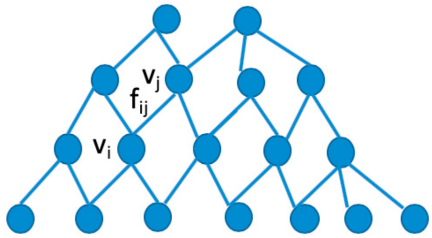

Fig. 3. Transmission Data flow when multiple parents exists

where $C(v_i)$ is the set of child clusters of v_i, and f_{ij} is the size of data received from child cluster v_j.

2. Flow constraints:

$$W_i = \sum_{vi \in P(v_i)} f_{ij} \tag{8}$$

where $P(v_i)$ is the set of parent clusters of v_i, and f_{ij} is the size of data transmitted from v_i to v_j.

3. Energy constraints:

$$E_i = e_i + \alpha \sum_{v_j \in C(v_i)} f_{ij} + \beta * W_i + e(W_i) \tag{9}$$

4. Objective function:

$$min\ max\ E_i \tag{10}$$

4. Simulation Result

In the proposed work I have implemented two algorithms to evaluate the network lifetime which is to be maximized. The network is deployed in such a way that the distribution of the nodes is uniform in the network. The network lifetime should be high which signifies that the time until the first node dies with respect to the number of rounds. The number of rounds signifies the transmission of the data unit to the base station from the source. The Simulation is run on different number of nodes such that, 100, 200,300, and 400 while the total area is 1000 x 1000 m. The aggregation energy is 5 nJ/bit while No of bit per second is 100. The simulation completed based on 400 rounds, initial energy is 0.5 J, and the message size is 4000 bits.

In the simulation the network is evaluated to the 400 rounds maximum and the performance is evaluated in terms of the network lifetime and latency. The network is run 3 times for each number of nodes, then the network lifetime is evaluated and the average network lifetime is noted down for each run. In this way, I have evaluated the performance in the network. For each round and transmission among the CHs the various evaluations are done such as Energy of node to transmit the packets for each round, Total energy of the CH for the particular round, Total aggregation energy etc. which are the important parameters to evaluate the performance of the network coz if energy consumption increases then there are lot of chances of the nodes to be dead which is responsible for the poor network lifetime. Also if the more energy is depleted then the overhead consumption also increases which increases the congestion in the network and the performance will also degraded. So this is the important evaluation which needs to take care for high network lifespan.

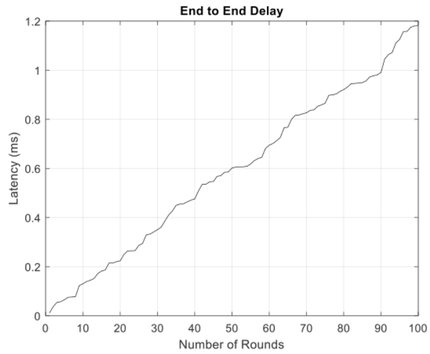

Fig. 4. Latency (ms)

Figure 4 shows the latency of the network which signifies the end to end delay and shows that the proposed approach is able to achieve low end to end delay from source to destination. If the latency in the network is low then the high data rates are there and high packet deliveries with low packet losses. If the latency is low then there will low overloads of the queues for the packets to be transmitted on the nodes then then there will less overhead consumption among the nodes also which leads to the less failures of the nodes and increases the lifespan of the network.

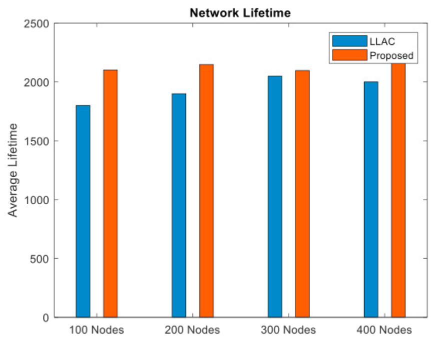

Fig. 5. Average Network Lifetime for ILP Algorithm

The figure 5 shows the average lifetime using 1st proposed algorithm. The above figure compare average lifetime after running simulations multiple times. The average lifetime must be high which signifies the network reliability and efficiency in the transmission of the packets from source to the destination without any node failures. The objective is to reduce the Data size constraint which will reduce the overhead in the network among the CH transmissions during data aggregation and also reduces the chances of the failures among the nodes and increases the lifespan of the network which is the desired objective.

Figure 6 shows the average lifetime evaluation using second algorithm approach which deals with the data flow reduces the data flow constraint for the high packet deliveries and low energy consumption. The main objective is to minimize the maximum energy consumption among the CH which will reduce the less node failures and increases the residual energies for the CH nodes to transmit the packets with more number of rounds. So Figure 6 shows the lifetime of the proposed work and shows that the proposed work is achieving high lifetime of the network which signifies the time until the first node dies.

5. Conclusion

I investigate the lifetime and security-aware clustering problem in WSNs and propose an ILP algorithm for constructing a lifetime and latency-aware, shortest path graph tree for routing topology. I have evaluated the approach

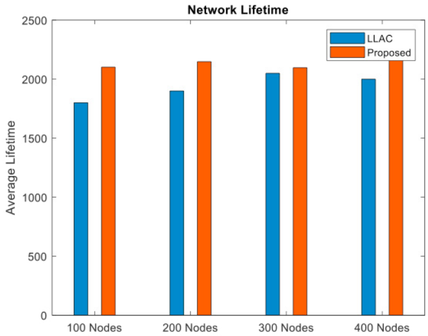

Fig. 6. Average Network Lifetime for LP Algorithm

using network instances with 100, 200, 300 and 400 nodes in the uniform distribution. The simulation results show that the approach significantly outperforms LLAC approach. To my knowledge, this is the First approach to this problem.

I assume that each cluster head performs perfect data aggregation. In some applications, the data sensed by each sensor node must be delivered to the base station. For those applications, different energy consumption model must be derived. Hence, our approach may not be efficient. I will investigate the clustering and routing tree construction problems for those applications in the future research. Furthermore, the assumption on perfect link quality made in this paper is not realistic. The transmission between two sensor nodes may not be always reliable. Another future research problem is to construct lifetime and latency-aware clusters and a routing tree by considering link quality.

References

[1] V. Rathod and M. Mehta. Security in wireless sensor network: a survey. Ganpat University Journal of Engineering and Technology, 1(1):35-44,2011.
[2] C. Wang, T. Feng, J. Kim, G. Wang, and W. Zhang. Catching packet droppers and modifiers in wireless sensor networks. In Sensor, Mesh and Ad Hoc Communications and Networks, 2009. SECON'09. 6th Annual IEEE Communications Society Conference on, volume 23, pages 835-843. IEEE,2012.
[3] S. P. Singh and S. Sharma. A survey on cluster based routing protocols in wireless sensor networks. Procedia computer science, 45:687-695, 2015.
[4] B. Bates, A. Keating, and R. Kinicki. Energy analysis of four wireless sensor network mac protocols. In Wireless and Pervasive Computing (ISWPC), 2011 6th International Symposium on, pages 1-6. IEEE, 2011.
[5] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In Proceedings of the 33rd annual Hawaii international conference on system sciences, pages 10-pp. IEEE, 2000.
[6] H. Lin, R. Xie, and L. Wei. Density, distance and energy based clustering algorithm for data aggregation in wireless sensor networks. In 2017 IEEE/CIC International Conference on Communications in China (ICCC), pages 1-5. IEEE, 2017.
[7] M. Mittal and K. Kumar. Data clustering in wireless sensor network implemented on self organization feature map (sofm) neural network" has been published in ieee international conference on computing communication and automation (iccca). SCOPUS indexed, 2016.
[8] X. Li, A. Liu, M. Xie, N. Xiong, Z. Zeng, and Z. Cai. Adaptive aggregation routing to reduce delay for multi-layer wireless sensor networks. Sensors, 18(4):1216, 2018.
[9] K. Maraiya, K. Kant, and N. Gupta. Wireless sensor network: a review on data aggregation. International Journal of Scientific Engineering Research, 2(4):1-6, 2011.
[10] B. Pourghebleh and N. J. Navimipour, "Data aggregation mechanisms in the internet of things: A systematic review of the literature and recommendations for future research," Journal of Network and Computer Applications, vol. 97, pp. 23–34, 2017.
[11] P. Singh, A. K. Gupta, and R. Singh, "Improved priority-based data aggregation congestion control protocol," Modern Physics Letters B, vol. 34, no. 02, p. 2050029, 2020.
[12] N. A. Humidi, V. Chowdhary et al., "Lightweight data transmission scheme based on data aggregation technique in wireless sensor networks," Girish, Lightweight Data Transmission Scheme Based on Data Aggregation Technique in Wireless Sensor Networks (May 17, 2019), 2019.
[13] R. Wan, N. Xiong, Q. Hu, H. Wang, and J. Shang, "Similarity-aware data aggregation using fuzzy c-means approach for wireless sensor networks," EURASIP Journal on Wireless Communications and Networking, vol. 2019, no. 1, p. 59, 2019.