Design and implementation of energy-efficient hybrid data aggregation in heterogeneous wireless sensor network

Mohamed Muthanna Al-Heeti¹, Jamal A. Hammad², Ahmed Shamil Mustafa³

¹Department of Electrical Engineering, College of Engineers, University of Anbar, Al-Ramadi, Iraq
²Department of Clinical Laboratory Sciences, College of Pharmacy, University of Anbar, Al-Ramadi, Iraq
³Department of Computer Engineering Techniques, Al-Maarif University College, Al-Ramadi, Iraq

ABSTRACT

Heterogeneous wireless sensor network (HWSN) is a trending technology in both the industrial and academic sectors, consisting of a large number of interconnected sensors. However, higher energy consumption and delay are significant drawbacks of this technology in applications such as military, healthcare, and industrial automation. The main objective of this research is to enhance the energy efficiency of HWSN using a clustering technique. In this article, a novel approach, namely power optimization and hybrid data aggregation (POHDA), is proposed to address these challenges in HWSN. POHDA-HWSN focuses on power optimization and congestion avoidance through effective CH selection using hybrid data aggregation based on parameters such as residual energy, distance, mobility, threshold value of the node, and latency. By weight-based effective cluster head (CH) selection, the energy consumption, end-to-end delay, and overhead during communication are reduced in this network. The POHDA-HWSN approach considers specific parameters to compare the results and outcomes with earlier research such as HCCS-WSN, FMCA-WSN, and APCC-WSN. The results prove that the proposed POHDA-HWSN approach achieves higher energy efficiency and delivery ratio.

Keywords:
Effective cluster head selection
Heterogeneous wireless sensor network
Hybrid data aggregation
Performance analyses
Power optimization

1. INTRODUCTION

A heterogeneous wireless sensor network (HWSN) is composed of a large number of sensor nodes that are distributed in a specific coverage area to collect data such as temperature and humidity levels. HWSN sensors have unique features, including wide coverage, battery levels, and functional behavior, and they are self-organizing, enabling collaboration with real-time operating systems. Recently, HWSN has become a highly active research area that attracts both industrial and academic attention. The applications of HWSN have expanded significantly, including real-time applications such as military, healthcare, and environmental monitoring, to ensure high-quality communication. The main advantage of HWSN is that the sensors used for data transmission are highly cost-effective and compact. However, the battery capacity of these sensors is limited, posing a significant challenge to achieving high energy efficiency while minimizing energy consumption during communication in the network. Therefore, energy optimization is one of the key solutions for addressing issues raised in earlier research on HWSN [1], [2].

The network architecture of HWSNs is designed with both static and mobile sensors. The deployment of these sensors is arranged in close proximity to enable easy access to neighboring sensors within their
coverage area. However, due to the dynamic nature of the network and its random deployment, the connectivity among the sensors is brittle. In many cases, the sensors are scattered in remote and hazardous areas where communication is challenging. The communication pattern is ubiquitous, with a large number of sensors interconnected through wireless connections. The primary challenge is to reduce energy consumption in this network model [3]. As a result of this drawback, communication in HWSNs is prone to faults and failures. Earlier research has focused on identifying where energy is most consumed. The conclusion was reached that the network’s energy is exhausted during the process of data aggregation and forwarding. Thus, energy-aware computing is still an open research area. The most commonly used technique for energy conservation in HWSNs is clustering [4]. Clustering involves a group of sensors electing a node to be the master node cluster head (CH) in their coverage area. The CH collects data from the nodes and distributes it to the respective destinations [5]–[10].

CH in HWSN play a crucial role in inter-cluster communication, intra-cluster communication, and communication with the base station (BS). To achieve effective data aggregation and forwarding, an optimal CH selection process is necessary to reduce unnecessary transmissions, thereby leading to reduced energy consumption [11]. The most challenging issues in clustering are effective CH selection and energy dissipation, and recent research in HWSN has focused on addressing these problems. In this research article, we propose a novel model that aims to optimize energy consumption and CH selection in HWSN networks. The primary contributions of this research are outlined below.

- The contribution of the research is improving the energy efficiency and lifetime of the HWSN network. To achieve that we developed a power optimization and hybrid data aggregation (POHDA) mainly concentrates on the reduction of energy consumption, delay and routing overhead.
- This approach provides effective CH selection based on weight oriented parameter calculations which assist in enhancing the lifetime. Once the CH is selected then the communication is performed in terms of inter cluster communication, intra cluster communication as well as CH to BS data transmission.
- Using power optimization process the efficiency is improved by performed perfect path selection for data transmission hence the network works with mobile nodes it becomes very essential. Implementation of the proposed POHDA approach is performed in NS2 and it is compared with the earlier approaches.

This research is organized as follows: section 2 elaborates the related researches about the clustering process in the sensor networks. Section 3 detailed about the proposed POHDA approach which includes power optimization and effective CH selection. Section 4 the evaluation of the network parameters is performed and the results are discussed. Finally, the conclusion and the future direction are given in the section 5.

2. RELATED WORKS

Azarhava and Niya [12] presented a novel approach to reduce power consumption in wireless sensor networks by utilizing the energy harvesting protocol to increase energy. However, the throughput of the network is decreased. Tyagi et al. [13] developed an internet of things (IoT) architecture for smart agriculture and improved it is energy efficiency by implementing back-propagation neural network (BPNN) and adaptive mutation particle swarm optimizer (APSO) neural network. Li et al. [14] presented an approach to increase the throughput by introducing a joint power allocation on multiple routing configuration (MRC) protocol. Zhang et al. [15] suggested an ensemble model to improve the spectral efficiency and power control of multiple-input and multiple-output (MIMO) networks. Li et al. [16] proposed a framework to reduce energy consumption. Du et al. [17] developed an algorithm to enhance the channel estimation efficiency by resolving the energy allocation problem. Least square channel and beamforming scheme is employed to optimize the channel estimation. Wang et al. [18] proposed a hybrid method of gradient descent (GD) and BS to reduce the computational complexity. Zhu et al. [19] presented an IS-k-means clustering algorithm to monitor the energy utilization. Ali et al. [20] developed a rank-based clustering algorithm to maximize the network lifetime by reducing the power consumption of the nodes. Umbreen et al. [21] developed a clustering model to minimize energy consumption by an effective CH selection process and avoid the overhead of cluster creation by re-clustering when necessary. Lipare et al. [22] proposed a method to improve the energy efficiency and lifespan of WSNs. This method utilized the fuzzy C-means algorithm for the formation of clusters, and the Sugeno fuzzy system is used to choose the cluster head, whereas to optimize the fuzzy rules.

Tao et al. [23] presented an unequal clustering algorithm based on fuzzy logic theory to increase the energy efficiency of the system by reducing power consumption of the sensors. Choi et al. [24] presented a model to enhance the transmission rate of the data. Lata et al. [25] developed a LEACH protocol with fuzzy logic based algorithm to increase the network lifetime. Ren and Yao [26] suggested an uneven clustering protocol to improve the network efficiency. Aydin et al. [27] developed two algorithms (GA and ANN),
to enhance the network energy utility. Wang and Zeng [28] proposed a novel approach to boost the energy efficiency of the network namely hierarchical clustering node collaborative scheduling method. To reduce the energy consumption of the network collaborative sensing is used node disks. Khedr et al. [29] developed an approach to reduce the data delivery latency of the network called as fuzzy C-means with routing protocol clustering model. Huang-Shui et al. [30] presented a new method to increase the lifespan of WSNs in an effective manner. For that purpose affinity propagation clustering approach and as well chaotic lion swarm optimization is used for optimal path selection in the network. This method produces moderate results in terms of throughput, energy consumption and lifetime but however it needs improvement. Based on the research study which is mentioned in this section. Once after analyzing the earlier approaches it is understood that the network needs further improvement in terms of energy efficiency and network lifetime. For that purpose, in our research power optimization with hybrid data aggregation is concentrated and it is elaborated in the upcoming sections.

3. PROPOSED POHDA APPROACH

The proposed POHDA approach is designed specifically to improve energy efficiency in HWSN. This section is divided into four parts: network model, energy model, power optimization, and hybrid data aggregation. Figure 1 shows the workflow of the proposed POHDA approach. Each of these parts plays a crucial role in achieving the overall goal of energy efficiency improvement in HWSN.

![Figure 1. Workflow of the proposed POHDA approach](image)

3.1. Hybrid cluster head selection

The process of data aggregation in the HWSN in an effective manner is extremely demanding. According to the objective, it is necessary to compile useful messages; the priority of the mission will also determine the quantity of messages needed. It was only necessary to gather information for missions like climate tracking, forest fire monitoring, and farming purposes. Which is involved in the major environmental changes. However, in operations like military applications, and robot location tracking, where the operation is critical, little modifications or ineffective messages may be significant, on that cases aggregation is very essential. Ant colony algorithms or some alternative optimization algorithms are an appropriate answer to the optimization problem with the data gathering process. However, the difficulty of the optimization method will have a detrimental effect on the sensor’s lifespan. The lifetime of the network can be improved using the classification algorithm but the computational capacity of the sensor never support classification problem which leads to increase of energy consumption in the network. Performance of data aggregation using clustering based on the geographical area is the best solution for the classification problems.

3.1.1. Concept of clustering

At the initial stage the each node perform neighbor discovery where it transmits a broadcast message to determine its neighbor inside its coverage area. Every sensor has two groups for storing the data of its neighbors: $NBR_1$ collects neighbor node ID and $NBR_2$ collect node ID, coordinates, and remaining energy of their neighbors. The primary criteria which are concentrated in the process of CH selection are the number of neighbor degree and the residual energy of the node. Approximately 40% of the initial energy is recognized as the energy threshold. The CH is chosen as the node with the highest neighbor degree and residual energy greater than its threshold. At the end of the CH selection process the CH broadcast a request message to all the nodes which are present in its coverage area. Then the process of data aggregation and route construction are carried out. Communication is carried out in terms of inter cluster communication and intra cluster communication in single hop manner. In some cases multiple hop communication is carried out if the base station is far away. To avoid collecting duplicate data, it additionally set the time value time1 with data transmission. To avoid communication with the dead node re-routing is performed in clustering process.
3.1.2. Process parameters of cluster head selection

In general using certain parameters in any network CH is selected to perform data aggregation and effective communication in it. In our CH selection process the parameters which are considered are residual energy, distance, mobility, node threshold value and latency. The calculation of those parameters for the purpose of CH selection is elaborated below. At the initial stage residual energy is calculated and it is defined as that every node estimates the amount of energy it has left for each round. Each node’s energy usage varies from that of the other nodes once they begin communicating and receiving data, which causes the nodes to eventually lose energy. The variation in energy depletion is proportional to their inter-communication or intra-communication distance and the mathematical expression for the calculation of residual energy is described in (1):

\[ P_1 = E_{\text{residual}} = \frac{E_{\text{remaining}}}{E_{\text{initial}}} \]  

(1)

In (4), the term \( E_{\text{residual}} \) implies the residual energy of the node, \( E_{\text{remaining}} \) implies the energy present in each sensor at the end of certain transmission and \( E_{\text{initial}} \) implies the initial energy of the node. The mechanism through which nodes lose energy is significantly influenced by distance. In comparison to remote nodes, closer nodes to the BS consume minimum energy to transfer data. The distance calculation is performed in the network with the help of the Euclidean formula. The mathematical expression for the calculation of distance between any two nodes is expressed in (2):

\[ P_2 = \text{Distance} = \sqrt{(A_y - A_{\text{new}})^2 - (B_y - B_{\text{new}})^2} \]  

(2)

In (2), the distance is measured between any two nodes with the coordinates of \((A_x, B_x)\) and \((A_{\text{new}}, B_{\text{new}})\). Hence it is HWSN the mobility of the sensors are also considered for the process of CH selection and here nodes can migrate in any direction and at any moment. There is a possibility of link failures due to high mobility in the network. Rapidly moving CH might allow the overall cluster to collapse, wasting time and resources. Additionally, frequent re-clustering consumes more energy. As a result, the mobility factor of sensors was taken into consideration. A random process is also used to choose the node’s new location. By determining the change between the node’s prior and new positions, the mobility level can be determined. The degree of mobility is determined using (3):

\[ P_3 = M_{\text{degree}} = \sqrt{(A_{\text{new}} - A_{\text{prior}})^2 + (B_{\text{new}} - B_{\text{prior}})^2} \]  

(3)

In (3), the terms \((A_{\text{new}}, B_{\text{new}})\) implies the coordinates of the node’s new position and the term \((A_{\text{prior}}, B_{\text{prior}})\) implies the coordinates of the prior position of the node. A node’s movement and possibilities to serve as a CH are inversely related, therefore a node with highly mobile has limited opportunities to be chosen as a CH. The threshold value determines the number of nodes which are present in the coverage area of the node or relative node count of the network. The mathematical expression for the calculation of the threshold value is described in (4):

\[ P_4 = Thresh_{\text{value}} = \sum_{i=0}^{n} L(i) \times Thresh_{\text{dis}} \]  

(4)

In (4), the term \( L(i) \) implies the number of neighbors present in the coverage area of the node, and \( Thresh_{\text{dis}} \) implies the threshold distance of the node. Finally, the latency of the node is collected from the node history and mathematically expressed in (5):

\[ P_5 = N_{\text{delay}} = \sum_{i=0}^{n} N_n \times N_{AL} \]  

(5)

In (5) considers two parameters for the process of CH selection in the network: \( N_n \), which represents the number of nodes present in the coverage area of each node, and \( N_{AL} \), which represents the average latency of the node. CH selection is a critical process in HWSNs as it can affect energy efficiency, network lifetime, and data accuracy. Thus, (5) provides a useful tool for selecting an appropriate CH by taking into account the number of nodes and the average latency of each node in the coverage area.

3.1.3. Effective cluster head selection

In our proposed approach, several parameters are considered for CH selection, including residual energy, distance, mobility, and packet delivery ratio. After establishing the conditions for CH selection, each
node’s energy is compared to a predetermined threshold, and the node with the highest weightage is considered the CH at the initial stage. The CH then uses a time-division multiple access (TDMA) schedule to perform communication between its cluster members with allocated time slots.

Furthermore, the CH performs data aggregation among the cluster members through inter-cluster communication, and it forwards the data to the BS using the inter-cluster communication process. The weighting of each node is determined using (6). By considering multiple parameters in the CH selection process, our proposed approach can effectively identify an appropriate CH that can optimize the performance of the network.

\[
Weightage = \frac{W_1 + P_1 + W_2 + P_2 + W_3 + P_3}{W_4 + P_4 + W_5 + P_5} \quad (6)
\]

From (6) the calculation for CH selection is performed in the network. Once after selecting the CH the communication is carried out using the inter cluster and intra cluster manner. Hence the nodes in the HWSN networks are mobile in nature some of the other activities such as the open leaving the cluster and node joining into new clusters are frequently happen. At the time the new node will send the request message to the new CH then joins the cluster otherwise all the CH will send periodical request message in flooding manner to get connect with the new nodes which are present in its coverage area. Any new node can accept the request and get joins into the new cluster. This is the process behind the proposed approach. Through this method communication becomes highly effective by reducing the energy consumption, delay and routing overhead of the network.

3.2. Power optimization

The primary goal of power optimization is to reduce the delay and overhead produced in the network. For that purpose in this HWSN network we use TDMA model to allocate the power and other resource to the nodes in the network. In general the TDMA model utilizes the time resource and it is separated into certain time slots. In each slot data transmission is performed through inter cluster and intra cluster communication at each instant of time. The process of power allocation is performed in the range of \((n + 1, 2n + 1)\) which perform one hop communication inside the cluster and two hop communication outside the cluster. In order to measure the valid resource elements for the nodes the total cardinality set is calculated for each node, and it is mathematically expressed in (7):

\[
C_{set}(node) = \sum_{i=n+1}^{2n+1} \frac{R_{element}}{R_{element} - V_{element}} \ast T(n_{iteration}, V_{element}) \quad (7)
\]

In (7) considers two terms: \(R_{element}\), which represents the resource elements taken into consideration for each node, and \(V_{element}\), which represents the validated resource from the initial resource allocation. By using (7), overhead can be eliminated by using only the valid elements, and the representation can be minimized using the minimum number of bits. This process is mathematically expressed in (8). By reducing the overhead and minimizing the representation, (8) can improve the efficiency and accuracy of resource allocation in wireless networks. This approach can be particularly beneficial in HWSNs, where energy efficiency and network performance are critical factors.

\[
I_{set}(node) = \log(C_{set}(node)) \quad (8)
\]

Based on the calculations presented above, the positions of the nodes play a crucial role in optimizing the network’s resources. Given that the network type is an HWSN, we also need to consider the dynamic nature of the nodes in the network. However, to minimize delay and overhead, the power optimization process prioritizes active nodes. This approach is mathematically expressed in (9). By considering the positions and activity levels of the nodes, (9) provides a powerful tool for optimizing power consumption in HWSNs. This approach can improve the network’s energy efficiency and extend it is operational lifetime, which are critical factors for many HWSN applications.

\[
Q_{power} = \sum_{i=1}^{n} C_{set}(node) \ast (1 - \alpha)^{n-1} \quad (9)
\]

In (9) includes the experimental constant \(\alpha\), which is a critical parameter for the power optimization process in the HWSN network. By using this equation, the network can prioritize active nodes, minimize delay and overhead, and improve its energy efficiency and operational lifetime. This approach is particularly useful in HWSNs, where energy efficiency is a critical factor for many applications.
4. PERFORMANCE ANALYSES

In this section, we compare the performance of the proposed POHDA approach with three existing and well-known algorithms: HCCS-WSN [23], FMCA-WSN [24], and APCC-WSN [25]. To evaluate the performance of POHDA-HWSN, we analyze specific parameters such as energy efficiency and packet delivery ratio. We used NS2 software to implement the proposed approach, and conducted simulations with varying numbers of nodes (ranging from 50 to 500) and a simulation run time of 300 seconds. Initially, the nodes were randomly distributed within the 2,000x2,000 m coverage area of the network.

4.1. Energy efficiency calculation

As shown in Figure 2, the performance analysis of the energy efficiency parameter for the earlier methods and it gets compared with the proposed power optimization and hybrid data aggregation wireless sensor network (POHDA-WSN) approach. The proposed scheme additional follows the idea of power optimization that mainly concentrates on reducing the energy consumption of the network. Hence the HWSN is the dynamically varying network both the sender and the receiver are in movement during the process of communication. To achieve effective communication optimized power utilization is essential for that purpose power optimization is employed in this network and it greatly helps to achieve better performance.

4.2. Packet delivery ratio calculation

As shown in Figure 3, it compared the successful packets transmission of the proposed POHDA-WSN with the earlier works. The POHDA-WSN greatly reduces the collision occurrence in the network that greatly helps to increase the packet delivery ratio of the network during the process of data transmission. In HWSN network each packet colliding possibility is high that creates more packet loss for that earlier methods like HCCS-WSN and FMCA-WSN. But in the POHDA-WSN with the presence of power optimization packet colliding is highly reduce that leads to achieve high packet delivery ratio during the process of data transmission in the network.

Figure 2. Energy efficiency of the proposed method
Figure 3. Packet delivery ratio calculation of the proposed method

4.3. Results and discussion

This section discusses the performance of all parameters considered for evaluation and highlights the performance improvement of the proposed POHDA approach. The analysis shows that the proposed POHDA achieves high energy efficiency and packet delivery ratio compared to earlier approaches. The energy efficiency achieved using the proposed POHDA-WSN method is 93.41%, while compared to previous strategies such as HCCS-WSN, FMCA-WSN, and APCC-WSN, it reaches up to 82.03%, 85.17%, and 89.16%, respectively. Therefore, the energy efficiency of the proposed POHDA-WSN method is 4% better than APCC-WSN and 8% better than FMCA-WSN, and 9% better than HCCS-WSN. The packet delivery ratio of the proposed POHDA-WSN method is 89.12%, while compared to previous strategies such as HCCS-WSN, FMCA-WSN, and APCC-WSN reach up to 84.26%, 86.29%, and 88.26%, respectively. Therefore, the packet delivery ratio of the proposed POHDA-WSN method is 1% better than APCC-WSN, 3% better than FMCA-WSN, and 5% better than HCCS-WSN. Overall, the performance of the proposed method is higher due to power optimization and hybrid data aggregation process.

5. CONCLUSION

To achieve high energy efficiency in HWSN, it is essential to prolong the network lifetime, especially if the network comprises mobile sensors. This work presents the POHDA-WSN approach in HWSN, designed to increase energy efficiency and packet delivery by combining an effective CH selection process and power optimization to save energy during data transmission. The network delay and routing
overhead are significantly reduced due to the re-clustering process, and power optimization provides proper energy and load balancing. Effective CH deployment manages long-distance communication and reduces network collisions, resulting in increased energy efficiency. The simulation carried out in NS2 compared the proposed POHDA approach with recent research such as HCCS-WSN, FMCA-WSN, and APCC-WSN. The analysis of the parameters such as energy efficiency and packet delivery ratio demonstrated that the proposed POHDA-WSN achieved 9% higher energy efficiency and 5% higher packet delivery ratio compared to the earlier approaches. As future work, drones can be incorporated into the proposed POHDA approach to further increase the network density.

ACKNOWLEDGEMENTS
Author thanks University of Anbar for support this work.

REFERENCES
Design and implementation of energy-efficient hybrid data aggregation in wireless sensor networks (Mohamed Muthanna Al-Heeti)

Mohamed Muthanna Al-Heeti was born in 1989 in Iraq, he obtained his master’s degree in communication systems Engineering at University of Portsmouth (United Kingdom) in 2016. After this, he worked as lecturer and Research in Ministry of Higher Education and Scientific Research, University of Anbar. Al-Heeti studied during his Master the Computer Engineering, mobile communication, line of site (LOS), microwave signals, wireless communication, digital communication, security, and satellite communication. He can be contacted at email: mohamed.muthanna@uoanbar.edu.iq.

Jamal A. Hammad received his master’s in computer science from the University of Anbar, Iraq in 2019. He is currently serving as a lecturer in the Department of Clinical Laboratory Sciences at the College of Pharmacy and the Department of Computer Science at the University of Anbar. He is highly interested in communication, computer engineering, and digital signal processing (DSP). He can be contacted at email: jamal.ali@uoanbar.edu.iq.

Ahmed Shamil Mustafa received his master’s in Communication and Computer Engineering from Universiti Kebangsaan Malaysia (UKM), Malaysia in 2015. He is currently serving as a lecturer in the Department of Computer Engineering Techniques at Al Maarif University College. He is highly interested in communication, computer engineering, VANETs, and digital signal processing (DSP). He can be contacted at email: ahmedshamil90@gmail.com.