Wireless Sensor Network: Energy Base Data Aggregation

Mr. Suraj Kumar Ram
Student, Amity University Chhattisgarh

Mr. Mohammed Bakhtawar Ahmed
Assistant Professor, Amity University Chhattisgarh

***Abstract*— Now-a-days, Wireless sensor networks play an important role in a wide range of applications. All the sensor nodes equipped with non-replaceable batteries for monitoring the desired region by fixed topology. By using WSNs everywhere, it becomes the main key of the big data in the emerging world. WSNs monitor physical phenomenon changes like temperature, pressure, ambient light, solar radiation and humidity consuming a large amount of energy. Energy consumption is a primary challenge in WSNs. To decrease the energy consumption by network, this paper proposes modified energy saving based privacy preserving data aggregation method to eliminate the redundancy among sensing data. This proposed method expands the system lifetime by reducing the number of transmission among the nodes and the sink. The efficiency of proposed method has been correlated with state-of-arts methods. The execution outcomes show that the proposed MEBPP data aggregation method send more number of packets with less energy utilization and more network existence.

Keywords— Wireless Sensor Network, Data Aggregation, Slicing, Mixing, Aggregator.

I. Introduction

Wireless sensor network (WSN) consists number of different self-organizing sensor hubs that perform detecting, processing and transmitting the data to the specified destination with low energy consumption [1]. WSN Routing is truly challengeable because of restricted assets like battery reinforcement, low information rate, less communication range and self-configurable. In certain situations, Sensor hub is outfitted with energy that is hard to supplant in brutal condition. Accordingly the network's lifetime relies on accessible charge in the battery of sensor hubs.

To improve the system lifetime, clustering between the nodes can be performed. The hierarchical routing protocols gives optimum energy efficiency. Rather than making group, sensor network ought to be isolated into numerous locales. In every locale, the sensor hubs communicate the data packet in an equal manner resulting in packets being sent without collision. This is a simple approach to aggregate the packets with less energy utilization. The data aggregation (DA) is registering a smaller portrayal for some of information that certainly represents to the enormous gathering. The method of aggregation of data is performed by a computing agent called an aggregator and usually performed at the Base Station (BS).

To reduce data flow, thus saving vitality, the process of aggregation may be distributed and networked. Processing can be done at all nodes or at designated nodes. WSNs are utilized to quantify ecological conditions such as contamination levels, temperature, sound, moistness, wind speed and bearing. At first WSNs are utilized for military activities yet its application has since been stretched out to wellbeing, traffic, and numerous different regions [2]. There exists a base station to which these hubs communicate information. As the power of the hub is limited and is provided by the battery, any significant distance correspondence won't be proficient and henceforth the lifetime of the hub won't be optimal. The fundamental thought of this work is to diminish the quantity of dead hubs and subsequently increment the average existence of lifetime.

This paper has been arranged in the following manner. Section 2 depicts analysis of the existing DA methods. Segment 3 depicts the detailed explanation of data aggregation. Section 4 depicts the proposed technique with the triple homogeneous cycle in detail. Section 5 gives the comparative analysis of MEBPP method with other existing methodologies. The conclusion is eventually briefed in the Section 5.

II. RELATED WORKS

A variety of experiments are in progress in the field of WSNs. The goal of any methodology is to improve the network existence by means of reducing power usage within the network. Multi-Hop scenario is a decent method to improve lifetime. In a multi-hop network, every hub sends the data to another hub and afterward transfers the data to the sink. Usually, the aggregator performs the data aggregation process and transfers the aggregated data to BS through sink node. Numerous applications utilize the hierarchical aggregation algorithm for the process. In [3], two-layered hierarchal aggregation scheme has been proposed to remove the redundant data. Since, this method efficiently increase the system existence, most of the existing aggregation algorithm depends on the hierarchical scheme [4].

In [5], PECDA scheme proposes secure data privacy by cost-effective encryption techniques to evacuate the redundant data effectively and reduce energy consumption. The study of [6] introduces the secret key based data aggregation technique to conserve electricity and secure energy effectively. The journal [7], proposes privacy-preserving, energy-saving data aggregation scheme to successfully decrease communication costs and to ensure privacy thereby securing useful information. It also warrants the balanced energy of the nodes prolong the existence of the system. The study [8] introduced a new energy efficient mechanism which would implement a super node for DA. It extracts data from CHs and aggregates in order to minimize the data redundancy prior to transfer to the sink. For making the network safe throughout data forwarding and aggregation, [9] proposes OSM-EFHE which has an added advantage of reducing the average energy utilization of nodes in the specified network with a minimum delay of 1.2ms. The WEMER protocol for data aggregation in WSNs has been built for in [10]. In this method, the cluster head sends information to the leader node. This node then forwards the data to the base station. Node.

Clustering based data aggregation methods are the most popular methods in making the WSNs energy efficient [11]. Numerous clustering based DA methods have been developed to increase quality of service and for reliable communication [12].The LEACH algorithm is one of the clustering algorithms. It has two phases: 1) Setup Phase and 2) Steady Phase. In the Setup Phase CHs are established on the basis of the probability value and all data transfer takes place in the Steady phase. In this strategy, the CHs are chosen arbitrarily and the node determines the cluster in which they belong to. The CH gets input from the hub and reduces the complete number of bits by aggregating them and successfully communicating the aggregated data to the sink. This algorithm works better than the other traditional protocols. In [13], cluster-based DA technique has been proposed to analyze big data in WSNs. In [14], proposed DA method for distributed sensor networks depends on IEEE 802.15.4 and accomplished a trade-off among energy utility and the likelihood of decision error. In [15] Energy efficient tree clustering method is proposed for better energy consumption and for long network existence. Likewise, the Internet of Things is firmly identified with distributed computing and big data analysis [16, 17]. Data aggregation is generally utilized in IOT, for example, savvy urban communities, contamination checking, woods fire avoidance, military applications, etc.

Further improvement is needed in clustering based data aggregation methods. Because of the short correspondence separation of WSNs, the long and linear spatial structure needs the regular use of multi-hop transmission, which poses two main problems. First, the amount of data is highly unbalanced. There are significant amount of data to be transferred and stored in the region closest to the sink node. This leads to unbalanced power utilization. Second, propagation time rises exponentially. Even though, the clustering algorithm is well proficient for long and linear WSNs, the DA in long and linear WSNs is still challengeable in terms of energy utilization. This proposed method does not perform the data aggregation operation directly. Instead of that, it will perform the slicing and mixing operation based on the energy level of the leaf node. After that, the aggregation tree will perform the DA. The execution results shows that the proposed method efficiently minimize the energy consumption.

III. DATA AGGREGATION

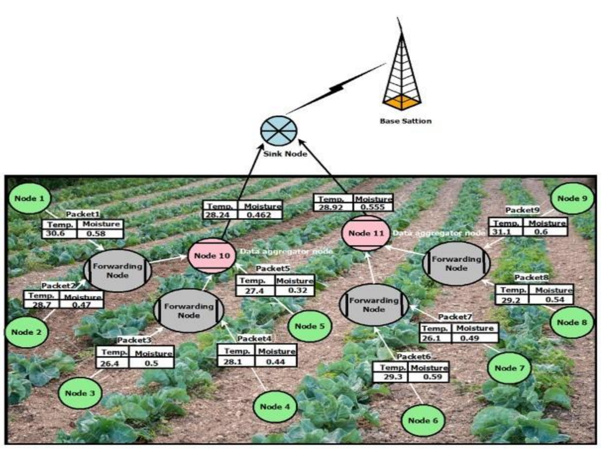

Figure: 1 Example of Data Aggregation

The method of aggregates collected information from the root node within the network system is referred to as data aggregation. As the stored data contains redundancy data, if it is passed to the sink node, the bandwidth will be lost. In addition, it expands the energy consumption of hubs [18, 19]. Subsequently, there is a prerequisite to minimize the energy utilization of nodes by decreasing the quantity of packets sent and in this way the network lifetime can be expanded. The key purpose of the DA is to eliminate repetitive data by extricating valuable data from the gathered data with the assistance of DA [20].

The figure 1 describes an example of data aggregation. In this network, numerous nodes are used to attain data including temperature and moisture content from a particular region. The information collected by the nodes are forwarded to their data aggregator node with assistance of sending hubs. The node 10 and 11 are the aggregator node which computes the average of gathered data then sends the same information to the sink Node. The nodes 10 and 11 are the aggregator nodes which computes and sends the average gathered data to the sink node. Thus, after DA, only 2-packets are sent instead of nine and these are received by the base station through the sink node. At node 10, the mean value of temperature is 28.24 ℃ and the moisture is $0.462 g/m^3$. Similarly, at node 11, the mean value of temperature is 28.92 ° C and the moisture is $0.555 g/m^3$. Thus, it tends to be reasoned that because of DA, the quantity of packets sent are less; as a result, data transmission is spared and less hub power is utilized to transmit the packets. This subsequently results in energy preservation in the hub / spare energy in the hub and there by expanding the network existence. In energy of the hub will be spared. Thus, it expands the network existence.

Significance of data aggregation

- It helps reduce network load and conserves energy.
- Aims to increase the power of the network.
- This also aims to minimize redundancy.

These significant factors result in an improvement in the latency for packet delivery across the system.

A. PROPOSED METHODOLOGY

The proposed data aggregation scheme has five phases: construction of aggregation tree, slicing, mixing, and aggregation.

3.1 Construction of Aggregation tree

The construction of aggregation tree is similar to HEEPP scheme. It consists of leaf nodes, forwarding nodes, sink node and base station. The leaf node gathers the information from atmosphere and sends to the forwarding node. . Forwarding node which is responsible for managing all the leaf nodes, accepts the data from leaf nodes and performs the initial stage of aggregation. On the other hand, as the number of forwarding nodes increases, the communication overhead will decrease. The forwarding node forward the data to the aggregator node where the data gets aggregated and is send to the BS through sink node.

Each node in this construction creates its own neighbor table and energy table. The energy table represents the remaining energy of the node and its neighbor node. Sensor hubs ought to have two boundaries set: NORMAL_CHILDREN and ADDITIONAL_NUM. *NORMAL_CHILDREN* is typically the highest number of children's nodes. In specific instances, nodes can utilize *ADDITIONAL_NUM,* which is the number of additional children nodes. If there is one intermediate node it is eliminated from the network, and its leaf nodes are supposed to look for a new intermediate node. Another ADDITIONAL NUM can be used by intermediate nodes to add additional leaf nodes [21]. The constructed aggregation tree is illustrated in figure 2. In this figure, the leaf nodes 1 to 9 slice and mix their pieces.

3.2 Slicing

Once the data slicing operations has been performed then the node encodes the data slices and transfers them to their CHs. Each node in the aggregation tree waits over a certain period of time to ensure that all the slices are sent. Then, each node decode the information utilizing a secret key shared only with the aggregator. This outcome shall be distributed to the parent.

So as to increase the convergence speed and reduce network traffic, nodes comply with the following guidelines when sending their data packets.

- The connectivity distance is one hop for all nodes. The leaf node transfers data slices to its neighbor node, which has the same CH.
- Similar data slice of the node must be transmitted once in a same time schedule.
- Node can transmit data to several hubs simultaneously, yet cannot execute the process of receiving and transmitting data at the same time, and the node cannot be received at the same time from different nodes.

The energy consumption of the leaf node gets reduced in such a way that the entire data are segmented into number of slices based on the Euclidean distance and the number of CHs. Consider M_i is the collected information of node *i and* m_{ij} is the size of the information packets transferred to node *j*. d_{ij} is the minimum Euclidean distance among node *i* and node *j*. Based on this distance d_{ij}, the leaf node *i* segments the collected data into *K* small data packets. The small data packets are represented as follows,

$$m_{ii} = \frac{1}{K} * M_i \quad (1)$$

The size of remaining (K-1) packets are represented as,

$$m_{i1}:m_{i2}:\cdots:m_{ij}:\cdots:m_{iK} = \frac{1}{d_{i1}}:\frac{1}{d_{i2}}:\cdots:\frac{1}{d_{ij}}\cdots:\frac{1}{d_{iK}} \quad (d_{i1} < d_{i2} < \cdots < d_{ij} < \cdots d_{iK}) \quad (2)$$

Consider unit distance is the smallest distance, the distance between other nodes can be written as,

$$d_{ij} = u_j * d_{i1}$$

(3)

Here, $u_j > 0$. Equation (2) can be re-written as,

$$m_{i1}:m_{i2}:\cdots:m_{ij}:\cdots:m_{iK} = 1:\frac{1}{u_2}:\cdots:\frac{1}{u_j}\cdots:\frac{1}{u_K}$$

(4)

According to the (1) and (4), we have

$$m_{ij} = \frac{1}{\sum_{j=1}^{K-1}\frac{1}{Ratio_{ij}}*Ratio_{ij}} * \left(\frac{K-1}{K}\right) * M_i$$

(5)

Equation (5) signifies the size of the data packet sent from the node. In this approach, the bigger packets are transmitted with nearest nodes, and the smaller packets are send to the neighboring nodes.

3.3 Mixing

After the slicing operation performed, the node encrypts the data packets and transfer to the AN through forwarding node. All the node in the aggregation tree wait for a while to ensure that all the slices are sent. Then the aggregator decrypt the received messages and aggregate all the received packets with the key shared with the aggregator. This consequences are transferred to the base station. Using figure 2. We can write the mixing operations of node. It can be represented as follows,

$$g_1 = d_{11}$$
$$g_2 = d_{22} + d_{12} + d_{32}$$
$$g_3 = d_{33} + d_{23} + d_{43}$$
$$g_4 = d_{44} + d_{34} + d_{64}$$
$$g_5 = d_{55}$$
$$g_6 = d_{66} + d_{46} + d_{76}$$
$$g_7 = d_{77} + d_{67}$$
$$g_8 = d_{88} + d_{78} + d_{98}$$
$$g_9 = d_{99} + d_{89}$$

3.4 Aggregation

When all data slices are acknowledged, each leaf node determines the total number of the slices and encodes the result with the mutual mystery key. The encrypted data is transmitted to the AN [22]. It collects all the data from its children nodes and aggregates it after decoding has been carried out. After the aggregation has been completed, the aggregation result will be transferred to its BS via sink node. The aggregation process has been depicted in figure 4. From the figure, the aggregated data can be written as follows,

$$G_1 = k_1 + g_1 + g_2 + g_3 + g_4 + g_5 + d_{1a} + d_{2a} + d_{3b} + d_{4b} + d_{5I}$$

$$G_2 = k_2 + g_6 + g_7 + g_8 + g_9 + d_{6c} + d_{7c} + d_{8e} + d_{9e}$$

$$K = G_1 + G_2$$

The modified EBPP scheme is explained as follows,

Modified EBPP Scheme

Step:1 Create the tree based aggregation structure.

Step:2 Set the time period for the nodes.

Step:3 Leaf hubs arbitrarily select the number of slices depending on the remaining energy.

Step: 4 Encode the slicing data and send them to neighbor hubs with one hop dependent on the energy table.

Step: 5 Decode the obtained data and merge the data.

Step: 6 Transmit the data to the intermediate nodes.

Step: 7 Intermediate nodes use the data validation technique to verify the consistency of the aggregate results.

Step: 8 If the intermediate nodes do not obtain a validation message, the data query procedure will be used.

Step: 9 Aggregate the received mixing info.

Step: 10 Give the output of the aggregation to the parent nodes before the base station.

The forwarding node has been waiting a long time to obtain information from them the children node with a data query mechanism. To reduce the waiting time, data authentication has been utilized. When the leaf node transfers the data packets completely to its parent node, it sends confirmation message and thus aborting the wait. When the intermediate node does not receive a confirmation response, the data query function will be used. This scheme increases the success rate of the data transmission and maintains of the accuracy of the data aggregation

IV. PERFORMANCE EVALUATION

In this section, the comparison of MEBPP data aggregation method with other existing DA algorithms has been performed based on various performance evaluation parameters. The existing methodologies such as EECS [23], OCCN [24], and EBPP [7] has been utilized for comparative analysis. The different parameters such as network lifetime, PDR and energy efficiency have been utilized. Each parameters are explained as follows,

Network Life time

It is the time period from the start of the network operation to the moment when the first sensor in the network runs out of energy.

the network lifetime correlation of different DA methodologies with the proposed MEBPP method. It is compared with EECS, OCCN and EBPP methods. This graph corroborates that the network lifetime of the proposed method is higher in terms of varying iteration levels compared to other existing DA methods.

Energy Efficiency

Each sensor in the network consumes a different amount of energy during data transmission. DA provides maximum performance with minimum vitality utilization in WSNs. It is a proportion of a measure of information effectively communicated in a sensor system to the total energy consumed to transmit those data.

$$Energy\ Efficiency = \sum_{i=1}^{n} \left(\frac{Amount\ of\ data\ successfully\ tr}{Total\ energy\ consumed}\right.$$

Figure 6 presents the comparison of energy utilization of the proposed method with existing data aggregation methods. From this graph, we can clearly observe that the MEBPP data aggregation method consumes less amount of energy when compared with other existing methods.

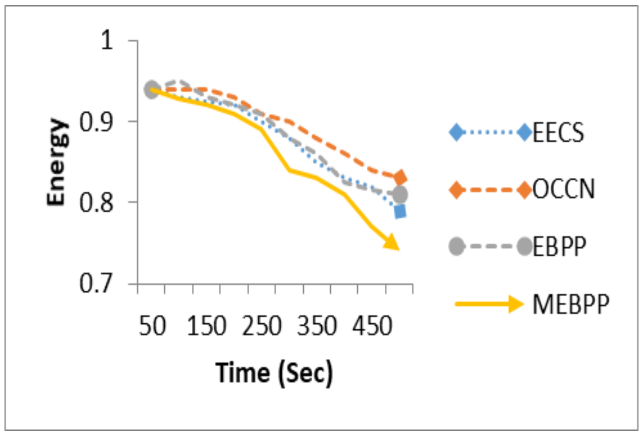

Figure: 2 Comparison of Energy level

Packet Delivery Ratio (PDR)

PDR is the proportion of successful transmission of information packets from sender to recipient hubs. Higher value of delivery ratio indicates a low loss in transmission. It is defined as follows,

$$PDR = \frac{Total\ number\ of\ packets\ received\ by\ the\ sink}{Total\ number\ of\ packets\ transmitted\ by\ all\ the\ nodes}$$

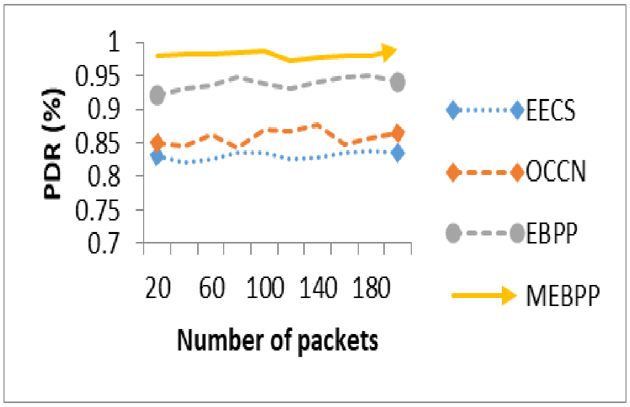

Figure: 3 Comparative analysis of Packet Delivery Ratio

The packet delivery ratio of MEBPP method is higher than the other existing methodologies. Since, in the MEBPP method the packet drop is lesser than the existing methodologies; it will automatically maximize the packet delivery ratio. This comparative analysis of PDR has been given in figure 7. From the above analysis, we can easily conclude that the proposed MEBPP data aggregation method gives better outcomes in terms of different evaluation parameters.

V. CONCLUSION

In order to reduce the energy consumption and to increase the network lifetime, modified energy saving based privacy preserving data aggregation method (MEBPP) has been proposed. For this, initially, we have constructed an aggregation tree. Each node in that aggregation tree segments its data into slices based on its remaining energy. At the same time, each leaf node send its data into neighboring nodes. The execution results have been compared with other existing data aggregation methods. From the analysis, we can conclude that, the MEBPP data aggregation method is well balancing in energy consumption, packet delivery ratio prolonging the network lifetime.

References

1. P.Paruthi Ilam Vazhuthi, Smart scheduling with parallel Data Aggregation protocol (SPDA) For Wireless Sensor Networks, Journal of Xi'an University of Architecture & Technology, Volume XII, Issue III, 2020, 5425-5430.
2. Li C., Liu Y.(2013) "ESMART: energy-efficient slice-mix-aggregate for wireless sensor network", International Journal of Distributed Sensor Networks.2013 (2): pp. 1–9
3. Tolani, M., & Singh, R. K. (2019). "Lifetime improvement of wireless sensor network by information sensitive aggregation method for railway condition monitoring". Ad Hoc Networks, 87, pp. 128–145
4. Guleria, K., & Verma, A. K. (2019)." Comprehensive review for energy efficient hierarchical routing protocols on wireless sensor networks". Wireless Networks, 25(3), pp. 1159–1183
5. T. Wang, X. Qin, Y. Ding, L. Liu, and Y. Luo, "Privacy-preserving and energy-efficient continuous data aggregation algorithm in wireless sensor networks," Wireless Personal Communications, vol. 98, no. 1, pp. 665-684, 2018.
6. J. Zhang, J. Zhu, Z. Jia, and X. Yan, "A secret confusion based energy-saving and privacy-preserving data aggregation algorithm," Chinese Journal of Electronics, vol. 26, no. 4, pp. 740-746, 2017.
7. Liming Zhou, and Yingzi Shan, "Privacy-Preserving, Energy-Saving Data Aggregation Scheme in Wireless Sensor Networks Journal of Information Processing Systems", Vol.16, No.1, pp. 83-95, February 2020.
8. Syed Gul Shah, Atiq Ahmed, Ihsan Ullah, and Waheed Noor, "A Novel Data Aggregation Scheme for Wireless Sensor Networks", International Journal of Advanced Computer

Science and Applications, Vol. 10, No. 2, 2019, pp. 585-590

9. M. Shobana, R. Sabitha, and S. Karthik," An enhanced soft computing-based formulation for secure data aggregation and efficient data processing in large-scale wireless sensor network", soft computing, 2020, pp. 1-12.
10. Abdulmalik Danmallam Bello, O.S. Lamba, "Energy Efficient for Data Aggregation in Wireless Sensor Networks, International Journal of Engineering Research & Technology (IJERT)", Vol. 9 Issue 01, January-2020, pp. 110-131
11. Afsar, M. M., & Tayarani-N, M. H. (2014). "Clustering in sensor networks: A literature survey. Journal of Network and Computer Applications", 46, pp. pp. 198–226.
12. Xu, L., Collier, R., & O'Hare, G. M. P. (2017). A survey of clustering techniques in WSNs and consideration of the challenges of applying such to 5G IOT scenarios. IEEE Internet of Things Journal, 4(5), pp. 1229–1249.
13. Din, S., Ahmad, A., Paul, A., et al. (2017). A cluster-based data fusion technique to analyze big data in wireless multi-sensor system. IEEE Access, 5, pp. 5069–5083.
14. Martalo, M., Buratti, C., Ferrari, G., et al. (2013). Clustered IEEE 802.15.4 sensor networks with data aggregation: Energy consumption and probability of error. IEEE Wireless Communications Letters, 2(1), pp. 70–73.
15. Chowdhury, S., & Giri, C. (2019). EETC: Energy efficient tree-clustering in delay constrained wireless sensor network. Wireless Personal Communications, 109, pp. 189–210.
16. C. Stergiou, K. E. Psannis, B. G. Kim, and B. Gupta, "Secure integration of IOT and cloud computing," Future Generation Computer Systems, vol. 78, pp. 964-975, 2018.
17. A. P. Plageras, K. E. Psannis, C. Stergiou, H. Wang, and B. B. Gupta, "Efficient IOT-based sensor big data collection: processing and analysis in smart buildings," Future Generation Computer Systems, vol. 82, pp.349-357, 2018.
18. S. Randhawa, S. Jain, Data aggregation in wireless sensor networks: previous research, current status and future directions, Wirel. Pers. Commun. 97 (3) (2017) pp. 3355–3425.
19. S. Boubiche, D.E. Boubiche, A. Bilami, H. Toral-Cruz, Big data challenges and data aggregation strategies in wireless sensor networks, IEEE Access 6 (2018) pp. 20558–20571.
20. P. Zhang, J. Wang, K. Guo, F. Wu, G. Min, Multi-functional secure data aggregation schemes for wsns, Ad Hoc Network. 69 (2018), pp. 86–99.