International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 11 (2019) pp. 2698-2703

Design and Implementation of Priority Hop Based Energy efficient Cluster Routing Algorithm for WSN Data Aggregation

M.Lakshmi[1], Dr P.Velmani[2], Dr P.Arockia Jansi Rani[3]

[1]Research Scholar (Part time Internal) Register No: 11964 Department of Computer Science & Engg, Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli-627012,Tamil Nadu, India.
[2]Assistant Professor of Computer Science The M.D.T Hindu College, Pettai, Tirunelveli, TamilNadu, India.
[3]Associate Professor of Computer Science & Engg , Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli-627012,Tamil Nadu, India
Corresponding Author

Abstract:

Wireless Sensor Network is an excellent technology that can sense, measure, and gather information from the real world based on some local decision process and it transmit the sensed data to the user. These networks allow the physical environment to be measured at any point and greatly increase the quality of the environment. For energy efficient transfer of sensed data, researchers proposed data aggregation methods, and hence data aggregation protocols are designed based on network architectures. It is divided into flat, hierarchal and location based protocols. The main objective of data aggregation is to increase the network lifetime by reducing the resource consumption of sensor nodes. Cluster based routing protocols plays vital role in energy consumption. Hence the main aim of this paper is design a new algorithm namely Priority Hop based Energy efficient Cluster Routing algorithm (PHECR) to evaluate the pollution level of air quality data. But this may be used to transform any type of aggregated pollution data.

Keywords: WSN, Routing Protocol, Data Aggregation

INTRODUCTION

The Wireless Sensor Network (WSN) is made up of large number of sensors and it consists hundreds or thousands of sensor nodes and among them at least one is a Base Station (BS). The sensor nodes are tiny devices which collect the information and transmit it to the BS. The basic components of sensors are sensing unit, processing unit, radio unit. The sensor nodes are collecting the data and transmit it to the Base Station which then sends this data to the user through the wireless communications [1].Several routing methods are available to pass information. Among them the cluster based routing is energy efficient method in which nodes those are having high energies are arbitrarily selected for processing and sending data while nodes those are having low energies are used for sensing information to the cluster heads. The cluster based routing protocols plays a pivotal role in achieving application specific goals. The cluster based routing protocols are classified into three broad categories such as block cluster based, grid based, and chain based routing protocol. [2] The objective of data aggregation is to reduce the required communication at various levels, and so as to reduce the total energy consumption. When energy consumption for aggregation is less than energy consumption for raw data transmission to the upper level, data aggregation saves energy. Eliminating the redundancy as well as energy consumption is always an issue which aggregation protocol considered [3]. Air pollution monitoring is considered as a very complex task but nevertheless it is very important. Traditionally data loggers were used to collect data periodically and this was very time consuming and quite expensive. The use of WSN can make air pollution monitoring less complex and more instantaneous readings can be obtained [4]. The monitoring systems are always hierarchical. The first layer is the monitoring networks. Sensor nodes sample environmental information and always transmit the data to the sink node in a wireless way. The second layer is the data uploading layer. The sink node uploads the data to the data center through wired networks or wireless ones. The data being transmitted to the data center is the raw data or the data preprocessed in advance by distributed nodes. The main differences of different proposed monitoring systems are in nodes and networks, such as different sensing platforms, different network topologies, different communication protocols, wired networks, or wireless ones. [5] Statistical methods have been used in almost every applied field to analyze the experimental data. Statistical methods are mathematical formulas, models and techniques that are used in statistical analysis of raw data. The application of statistical methods extracts information from data provides different ways to assess the outputs. Robust statistics provides an alternative approach to classical statistics methods [6]. The paper is organized as follows Section 1 describes Application Scenario Section 2 presents the Literature survey regarding the importance of cluster routing protocol, Section 3 outlines accurate air quality analysis levels, Section 4 Illustrates the proposed methodology, Section 5 Discusses simulation study and Section 6 Concludes the paper.

1. Application scenario

Now-a-days more modern technologies are available to monitor air pollution data in large scale industries in State and Central Pollution Control Board. But still there is a need to improve reporting mechanism. Air pollution is causing hundreds of millions of people in India to lead shorter and sicker lives

To help improve India’s air quality, researchers from the University of Chicago and Harvard Kennedy School have laid out five key evidence-based Policy recommendations in a new report recently.”A Roadmap Towards Cleaning India’s Air”

India is searching for the best way to balance the dual and, at times, conflicting goals of achieving economic growth while maintaining a clean environment. EPIC-India researchers have developed the Air Quality Life Index (AQLI), a metric that provides a means to predict the overall reduction in life expectancy caused by living in places with high levels of air pollution. These health costs are not restricted to a few urban areas. The AQLI is generated by combining global datasets on air pollution with published scientific evidence on the causal effects of pollution on life expectancy based on a natural experiment. There are five key policy recommendations.

Improve emissions monitoring by better aligning incentives of auditors.

Provide regulators with real time data on polluters' emissions.

Provide the public with information about polluters.

Use monetary charges for excess emissions.

Use markets to reduce abatement costs and pollution.

These include increasing the quality of information about polluters' emissions that is provided to both regulators and the public, as well as introducing market based instruments to deliver the largest impact. High levels of air pollution were once common place in developed nations such as the U.S., England, and Japan. These countries were able to address this problem by tightening regulations and introducing new policies. Today, India faces a similar opportunity to utilize a variety of tools to improve environmental quality. [7] one of the recommendations includes improving emissions monitoring by better aligning incentives of auditors. Hence it is evident that real-time reporting data on polluters, emission is necessary to take further action. [8]

LITERATURE SURVEY

Routing in WSN differs from conventional routing in fixed networks in various ways. There is no infrastructure, wireless links are unreliable, sensor node may fail, and routing protocols have to meet strict energy saving requirements. Many routing algorithms were developed for wireless networks in general [9]

Among them DSDV and LEACH are frequently used protocols by the researchers[10] DSDV is a proactive routing protocol and it is somewhat same as the conventional Routing Information Protocol (RIP) and has the only difference of having additional attribute in the routing table that is the sequence number. At each node of the network the routing information which is used while routing is stored using a table known as routing table. Routing table has the attributes; all the available destinations, the sequence number assigned by the destination node and the number of hops that is needed to reach the destination node and with the help of this table, communication between nodes in the network take place. Consistency among the routing table in the nodes is maintained by broadcasting regularly the routing information stored in the routing table to every neighbor. The broadcasted routing information contains the fields; the nodes' new sequence number, the IP address of the destination, the new sequence number assigned by the destination and the number of hops required to reach that destination. And the latest destination sequence number is used for making decisions to forward the information again or not. This latest sequence number is also updated to all the nodes which are passed by the information while transmitting within the network. [11]

LEACH is a self-organizing, adaptive clustering protocol that uses randomization to distribute the energy load evenly among the sensors in the network. In LEACH, the nodes organize themselves into local clusters, with one node acting as the local base station or cluster-head. If the cluster heads were chosen a priori and fixed throughout the system lifetime, as in conventional clustering algorithms, it is easy to see that the unlucky sensors chosen to be cluster-heads would die quickly, ending the useful lifetime of all nodes belonging to those clusters. Thus LEACH includes randomized rotation of the high-energy cluster-head position such that it rotates among the various sensors in order to not drain the battery of a single sensor. In addition, LEACH performs local data fusion to "compress" the amount of data being sent from the clusters to the base station, further reducing energy dissipation and enhancing system lifetime.[12]

The successor of the LEACH protocol that is used in data aggregation is W-LEACH, IB-LEACH and DAO-LEACH. W-LEACH can handle sudden changes in the underlying data that can happen according to the occurrence of events in the sensed field. It highlights the area of interest so that the sensors in the event area are more likely chosen to send data to their cluster heads.[13] The operation process of IB-LEACH consists of several rounds and each round is split into three phases: set-up, pre-steady and the steady state. In the pre-steady state phase, sensor nodes of a cluster are divided into three categories. CH, sensing nodes and aggregator Sensing nodes sense the environment and send sensed data to the aggregators. The aggregators aggregate the received data and send it to the BS. This reduces the energy consumptions of Cluster Heads. [14] The entire process of DAO-LEACH is divided into four stages: node deployment, cluster formation, optimal numbers of Cluster Head selection and aggregation via data ensemble. The optimal deployment of nodes and Cluster Head selection improves the load balance of the network. [15]

3. Accurate air quality analysis

Air Quality Index Summary Report displays an annual summary of Air Quality Index (AQI) values for countries or Core Based Statistical Areas (CBSA). To compute accurate air quality, AQI with all pollutant in a geographic area is to be measured.

Table 1: Air Quality Index Value Category

AQI CATEGORY	RANGE
0 to <50	Good
50 to <100	Moderate
100 to <150	UnHealtyforSensitive Group
150 to <200	UnHealthy
200 to <300	Very UnHealthy

Table (1) shows the AQI category and the corresponding air quality values in , parts per million (ppm) and parts per billion (ppb).

The point to be considered here is, the proper aggregation method used in industrial pollution monitoring. Mostly industries are sending average of air pollution data to the state Pollution Control Board. (Tamil Nadu Pollution Control Board).This may lead to the false analysis.

4. Proposed Method

This paper is a design of a new algorithm namely Priority Hop Based Energy Efficient Clustering Routing algorithm. In this algorithm sensor node which has higher energy is selected as cluster head for each hop level and sense the air quality index data in priority basis. The proposed method is tested in sensor network of varying size that is, number of nodes such as 30,40,50,55,60,65,70. The nodes are randomly deployed.

Base station is allocated automatically and all other nodes are found to which node neighbor for that. Maximum number of hop is assigned and number of clusters are formed and Cluster heads are formed based on the highest energy level and cluster members are sense the air quality index values and send to cluster head .Cluster head send air quality index value to base station finally base station reports polluted area in priority basis.
The algorithm is tested by creating backend coding in two files

1. hop.cc for maximum number of hop process and
2. mleach.cc for assigning energy to sensor node from that higher energy node is treated as cluster head

This algorithm is designed by combining the features of DSDV and LEACH protocol that is necessary for efficient data reporting of air pollution data. Fig (1) Shows Diagrammatic Representation of PHECR. Fig (4) shows the proposed algorithm.

5. Simulation study

For data aggregation the method available in the literature are use of various aggregation function and various Probability distribution methods. Mostly the aggregation function available in the literature are the classical statistical methods such as average and total functions that are affected by extreme values and may distort and reduce its usefulness.[16][17] So for this reason it is decided to use Robust mean aggregation in this proposed algorithm for better and useful data reporting. Robust mean such as Trimmed mean and Winsorized mean provides an alternative approach to classical statistics method. Trimmed mean is a statistical measure of central tendency much like the mean and median. It involves the calculation of the mean after discarding given parts of a sample at the beginning and the end of the whole data, and typically discarding an equal amount of both one. For most statistical applications, 5 to 25 percent of the ends are discarded. The trimmed mean is a useful estimator because it is less sensitive to outliers than the mean. In this regard it is referred to as a robust estimator. The winsorized mean is a useful estimator because it is less sensitive to outliers than the mean but will still give a reasonable estimate of central tendency or mean for almost all statistical models [18][19]. So this study was carried out aggregation function using Mean and Robust Mean to calculate air quality index data.

Hop Process Calculation
If nn<minimum coverage
nn in hop level 1
Else if
Hop level = hop count + 1
nn in hop level 2
Else
Hop level = hop count+1
nn in hop level 3

Formulas to find number of cluster head
TCH = NN * PCH / 100
Where,
TCH Total number of Cluster Head
NN Number of nodes
PCH Percentage of Cluster Head Allocation

Base Station Allocation
If $ i != $ j then
$ i is Base station
Else
$ i is neighbor for $j

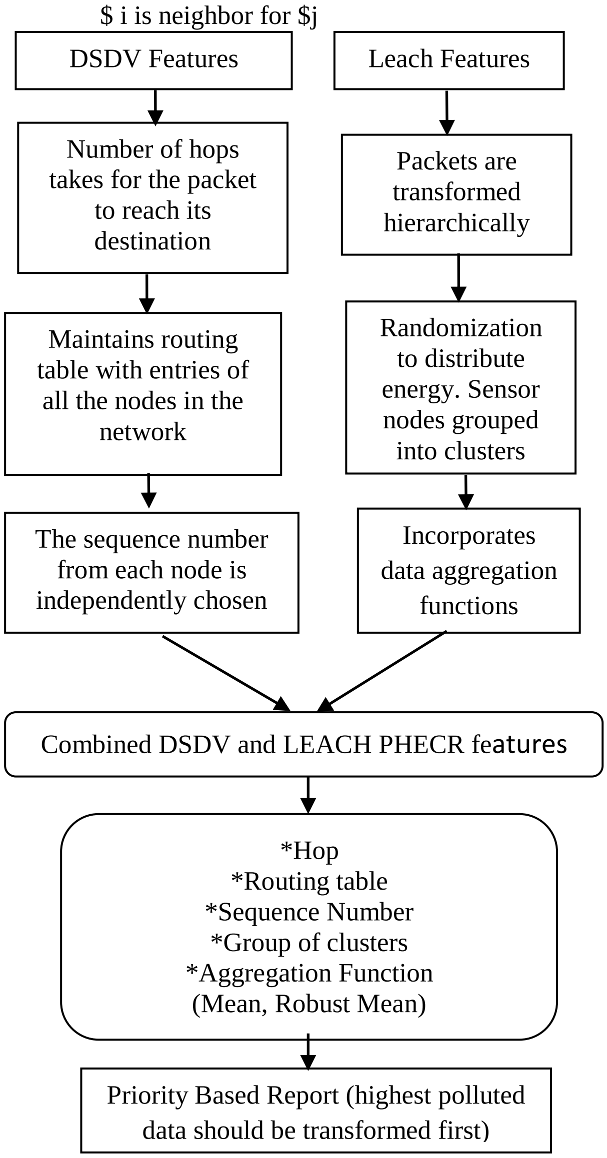

Fig 1 Flow Diagram of PHECR

```
Input :       Number of nodes, neib index,
 neiblist ,  aqi values
Output: Number of hop, Number of clusters, Polluted
range of AQI values
Algorithm: PHECR
1. N=n n; Number of nodes
2. for each I ,j ε  n n do
3      if i ≠  j then
4.          j <= coverage then
5.                      total no of neighbour for i
6.          end if
7. end for
8. for each nn ε  i do
9.   If neighindex(i)<neighindex(i+1) then
10.         hopnumber=min(neighindex)
11.         else
12.                     hopnumber=hopnumber+1
13.         end if
14. end for
15. for each  nn ε  hopnumber(i)
16.         If hopnumber(i)=nn(maxenergy(nn))
17.                     nn=ch
18.         Else
19.                     nn=cm
20.         end if
21. end for
22. If ch(i)>=max(aqi level)
23.         Print aqi data bast in time
24.End if
```

Fig 2 Algorithm PHECR

Explanation of above algorithm

Number of sensor nodes is given as input .Belong to that number of hops is assigned. Each Sensor node identify neighbor node for that. Sensor node need not to check itself for neighbor node. Each hop level contains at least one Cluster Head even if the minimum of sensor node is allocated for that hop to sense the air quality data. First hop Level is assigned due to the minimum coverage area. After that hop level is increased by 1 and the process should be continue until the maximum coverage area to be reached.

Each hop level allocated energy for all the sensor nodes which sensor node higher energy that sensor node is assigned as Cluster Head remaining sensor nodes are Cluster Members. Sensor node which has higher polluted range reaches the base station in time directly through the sensor node belong to the cluster head need not pass on any other cluster head. Finally base station report the polluted level of each cluster head.

Table 2. Parameters considerd in this study

S. No	Parameters Description	Value
1	Network dimension	600 m x 600 m
2	Number of nodes	30,40,50,55,60, 65,70
3	Simulation duration	3600 s
4	MAC	802.11
5	Number of cluster Head Location	Random
6	Routing Protocol	DSDV
7	Traffic source	CBR(Constant Bit Rate)

Table 3. Values of Mean and Robust Mean

No of nodes	55	60	70
Cluster Head Location	33	4	28
Mean	191.46 (UH)	148 (UFS)	196.06 (UH)
Trimmed Mean 20%	208.44 (VUH)	154.54 (UH)	214.88 (VUH)
Winzorised Mean 20%	203.84 (VUH)	147.82 (UFS)	199.33 (UH)
Trimmed Mean 25%	212.71 (VUH)	158.77 (UH)	226.0 (VUH)
Winzorised Mean 25%	208.46 (VUH)	150.64 (UH)	203.86 (VUH)

Table 3 shows the comparision of Mean, Trimmed Mean and Winsorized Mean(20%,25%) with the sensor network of varying size such as 30,40,50,55,60,65& 70 .For sample the table listed for aggregated value in three clusters by using Mean and Robust Mean. From that it is evident that Robust Mean gives better analysis, to report more polluted data.

Figure 3 shows the comparision of pollution category using estimates Mean, Trimmed mean 25% and Winsorized mean 20%. .Only after 25% elimination of data, the pollution category falls on similar type or some higher category of pollution range.

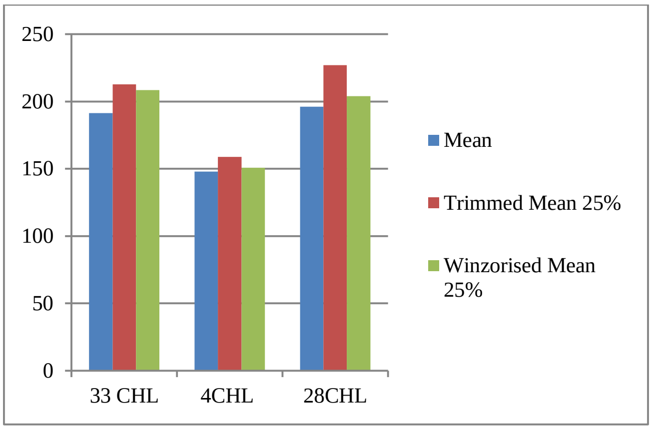

Fig 3 Graphical Representation of Table 3

Table 4. Priority Based Values

CLUSTER HEAD LOCATION	AQI DATA	TIME
23	290	774.30
17	267	775.40
17	258	776.50
19	256	777.60
17	242	778.70
19	239	779.80
19	215	780.90
17	198	782.00
19	195	783.10
17	193	784.20
17	190	785.30
17	185	786.40
17	179	787.50
17	145	788.60
17	142	789.70
19	133	790.80
23	91	791.90
23	80	793.00
23	68	794.10
23	58	795.20
17	50	796.30
23	46	797.40
17	31	798.50
17	27	799.60
19	6	800.70
17	2	801.80

Table 4 shows the data received by different cluster heads .The cluster head having maximum polluted data is forwarded first to the BS on the priority basis

6. CONCLUSION

To report most polluted aggregation air quality data, the proposed algorithm namely PHECR combines the features of DSDV and LEACH .The proposed algorithm is designed and implemented in certain WSN setup using the best aggregation function Robust Mean. It is concluded that for usable data reporting mechanism and to get real picture in pollution data reporting use of robust Mean is a best choice.

REFERENCES

[1] Vaishali P. Bawage,Deepak C. Mehetre "Energy Efficient Secured routing Model for Wireless Sensor Networks"*Proc. of International Conf* on Automatic and Dynamic Optimization Techniques (ICACDOT). International Institute of Information Technology, Pane 2016

[2] Santar Pal Singh [a],S.C.Sharma [b].,A Survey on cluster Based routing Protocols in Wireless Sensor Networks"International *Proc. of International Conf* on Advanced Computing Technologies and Applications(ICACTA-2015)Procedia Computer Science45(2015)687-695

[3] Geetika Dhand a ,S.S.Tyagi b*," Data aggregation techniques in WSN:Survey" *Proc. of 2nd International Conf* on Intelligent Computing, Communication & Convergence (ICCC-2016) Procedia Computer Science 92 (2016) 378 – 384

[4] Khedo,K.K., A Wireless Sensor Network Air Pollution Monitoring System.International Journal of wireless and Mobile Networks,May 2010.No.2(no. 2)

[5] Xu_Luo[1] and Jun_Yang[2] ,"A Survey on Pollution Monitoring Using Sensor Networks in Environment Protection", Journal of Sensors Volume 2019,

[6] Frederick J Gravetter ,Larry B.allnau Statistics for the behavioral science" .

[7] https://epic.uchicago.in

[8] www.ians.in

[9] Shio Kumar Singh, M P Singh, and D K Singh "Routing protocol in wsn- A Survey" IJCSES Vol.1,No.2,November 2010

[10] IBRIHICH OUAFAA, 1 ESGHIR MUSTAPHA, 2 KRIT SALAH-DDINE, 1 EL HAJJI SAID ,"Performance analysis of sleach, leach and dsdv protocols for wireless sensor networks (WSN)"Journal of Theoretical and Applied Information Technology 31st December 2016. Vol.94. No.2 © 2005 - 2016 JATIT & LLS. ISSN: 1992-8645

[11] Teressa Longjam* and Neha Bagoria** "Comparative Study of Destination Sequenced Distance Vector and Ad-hoc On-demand Distance Vector Routing Protocol of Mobile Ad-hoc Network", International Journal of Scientific and Research Publications, Volume 3, Issue 2, February 2013 1 ISSN 2250-3153]

[12] 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 Siences (HICSS '33), p. 223, January 2000.

[13] HanadyM. Abdulsalam and Bader A. Ali "W-LEACH Based Dynamic Adaptive Data Aggregation Algorithm for Wireless Sensor Networks" Hindawi Publishing Corporation International Journal of Distributed Sensor Networks Volume 2013, Article ID 289527, 11 pages http://dx.doi.org/10.1155/2013/289527

[14] A. Salim, W. Osamy, and A. M. Khedr, ``IBLEACH: Intra-balanced LEACH protocol for wireless sensor

networks," Wireless Netw., vol. 20, no. 6, pp. 1515_1525, 2014.

[15] A.G. Seminathan and S.Karthik,"DAO-LEACH: An approach for energy efficient routing based on data aggregation and optimal clustering in wsn",Life Sci J.,vol. 10 no.7,pp.380-389.2013

[16] M.Lakshmi[1] , P.Velmani, PhD[2], P.Arockiya Jansi Rani, PhD[3],"Study on Data Representation and Aggregation in WSN" International Journal of Computer Applications (0975 – 8887) Volume 165 – No.11, May 2017

[17] Nandini. S. Patil, Prof. P. R. Patil,"Data Aggregation in Wireless Sensor Network" 2010 IEEE International Conference on Computational Intelligence and Computing Research

[18] M.Lakshmi[1] , P.Velmani, PhD[2], P.Arockiya Jansi Rani, PhD[3]," Data Aggregation in WSN Using Robust Mean" International Journal of Research in Advent Technology, Vol.6, No.10, October 2018 E-ISSN: 2321-9637

[19] M.Lakshmi[1] , Dr P.Velmani[2], Dr P.Arockia Jansi Rani[3],"Comparison of Continuous Probability Distribution and Robust Mean in WSN Data Aggregation" IOSR Journal of Engineering (IOSRJEN) ISSN (e): 2250-3021, ISSN (p): 2278-8719 Vol. 09, Issue 1 (January. 2019), ||V (I) || PP 33-37