A Framework to Increase the Lifetime of a Sensor Network by Dynamic Cluster Head Selection and Re-Clustering

Vivek Yadav, M. Pushpalatha

Abstract -- In spatial correlated sensor networks, nearby sensor nodes show a high degree of similarity in their readings. Keeping this in mind, the monitored region is divided into clusters and one cluster head is assigned to a cluster which does the sensing work while other nodes are in sleep state. This results in considerable energy savings as only one cluster head representing other nodes in the cluster does the sensing and transmitting work. In this paper, a framework is proposed to add functionality to existing networks so as to increase the networks lifetime and also aid in better tracking of unusual phenomena such as forest fires. Increasing the lifetime has been done by the use of dynamic multiple cluster heads and tracking of unusual phenomena has been done on the basis of outlier detection.

Index Terms-- Lifetime Optimization, Spatial Correlation, Dynamic Cluster Head Selection.

I.INTRODUCTION

A wireless sensor network is a sensor network comprising of multiple sensor nodes which communicate with or via each other to send sensed information to an end station. They have many applications such as environment monitoring, structure monitoring, home automation, surveillance etc. Communication occurs instantaneously between nodes within communication range [10].

Sensor nodes can be manually or randomly deployed. In home automation or structure monitoring, sensor nodes are usually deployed in fixed positions. In unreachable areas such as environment monitoring areas, the nodes are randomly deployed, for example in case of flood hit areas.

Sensor networks are energy constrained networks. Each sensor node has limited sensor power to operate in its lifetime. Although new techniques have been developed to repower sensor nodes, still energy usage optimization remains a strong research area.

In case of randomly deployed networks, energy conservation is an important factor. One way to conserve energy is in the form of in network aggregation. In this technique the sensed data of many sensor nodes is accumulated, processed and send to the base station via only one node instead of all the sensing nodes that participated. The sensing nodes within a cluster select a cluster head according to some parameters such as distance, energy levels etc.

Manuscript received April, 2014

Vivek Yadav, Department of Computer Science and Engineering, SRM University, Chennai, India.

Dr. M. Pushpalatha, Professor, Department of Computer Science and Engineering , SRM University, Chennai, India.

Many cluster based approaches for data aggregation have been proposed in recent times. Most of these approaches form clusters in sensor networks where nodes in a cluster send their sensed information to their cluster head which in turn does the processing and sends the information to the end station (sink). The method to elect cluster head depends on different techniques.

Low Energy Adaptive Clustering Hierarchy (LEACH) [6] is a protocol that elects cluster heads keeping in mind the energy load distribution. It does so randomly by rotation. Each node chooses a random number, and if number is less than a predefined threshold, the node becomes a cluster head. In this way every node has a chance to become a CH depending on its residual energy. In HEED [13], cluster head is chosen based on a hybrid of its residual energy and communication costs. In PEGASIS [1], nodes form a chain with nearby sensor nodes to send and receive data. The gathered data moves from node to node, gets processed and eventually is sent to the base station by a designated node. In order to balance the overhead involved in communication between the designated node (head node) and the base station, each node in the chain takes turn to be the leader (head). In [2] each node is assigned a unique ID number and the node with the highest ID number in the cluster becomes the cluster head. In [3] a set of representative nodes are determined with high energy levels and wide data coverage ranges and selected as cluster heads covering a subset of area. Ghalib et al. [9] have given a technique in which only the cluster heads are responsible for exploiting spatial correlation of their member nodes and selecting the appropriate member nodes to remain active. The correlation is based on residual energy of member nodes. Each cluster head divides its clustered region into correlation regions and selects a representative node in each correlation region which is closer to the center of correlation region and has the higher residual energy. Hence, the whole field is represented by a subset of active nodes which perform the sensing task well.

Most of the above algorithms use a measure of residual energy to get selected as cluster heads. Another factor that affects random deployment in sensor networks is spatial correlation. Spatial correlation in sensor networks states that nearby sensor nodes monitoring an environment typically sense similar values. In randomly deployed sensor networks, the sensor nodes might get positioned very nearby. This results in waste of energy as many nodes will be monitoring the same area which can be monitored by only one node. To overcome this problem, research has been done to exploit spatial correlation in sensor networks. The spatial correlation increases with decreasing distance. Spatial correlation has been extensively used in formation of clusters in sensor networks. The proposed framework in this paper adds functionality to aggregated clustered

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

networks formed on the basis of spatial correlation to increase the network's lifetime.

The remainder of the paper is organized as follows: In Section 2, works related to spatial correlation are described. In Section 3, the new framework to increase the network lifetime is proposed and described. Section 4 describes the metric to measure the lifetime. Section 5 describes the expected outcome of the proposed scheme. Finally, Section 6 concludes the paper.

II.RELATED WORK

Spatial correlation has been exploited in many research works. Some works have used spatial correlation as a measure of distance and others as a measure of readings.

Yoon et al. [2] have given a technique in which clusters are formed based on spatial correlation. Only sensor values from the cluster heads is used to compute the aggregate and send it to the sink. This method takes into consideration a user provided threshold value and a node only joins the cluster if its value and cluster head value satisfy the formula $MR < CR \pm CR \times \tau$ Where CR (Cluster Head sensor reading), MR (Local sensor reading) and τ is the user defined threshold.

Ma et al. [4] have proposed an α -local spatial clustering algorithm for sensor networks based on the dominating set theory. By measuring the spatial correlation between data sampled by different sensors, the algorithm constructs a dominating set of the cluster head's as the sensor network backbone which is used to realize the data aggregation. In [8] on similar terms, the authors have proposed the data density correlation degree. The proposed correlation degree is a spatial correlation measurement that measures the correlation between neighboring sensor node's data.

[11] finds out the spatial correlation between sensor nodes based on Pearson coefficient which is then used to form clusters.

Chu et al. [12] exploit spatial correlation for approximate data collection, where probabilistic model is used to predict sensor readings. If the readings are accurately predicted, the sensor nodes do not send their readings to the sink, thereby reducing communication costs.

Shakya et al. [7] have proposed a basic spatial correlation function to model the correlation characteristics of the event information observed by multiple omni-directional sensor nodes. It takes into account spatial correlation based on distance not on measured readings.

The existing systems above form spatially correlated networks and select cluster heads to represent the respective clusters. Due to only a subset of nodes being activated at a time energy savings are considerable which helps increase the lifetime of the network. The existing systems only consider the use of one cluster head at a time for a cluster which can act as a limitation. Although this can lead to most accurate results, it results in overload on the single cluster head. This can deplete the energy of the node quickly and the network can die down.

Therefore in this paper a framework is proposed to add multiple dynamic cluster heads to the existing network. The key point to note is that it is not compulsory for a cluster to have multiple cluster heads. How many cluster heads are there in the network depends on the node satisfying a minimum threshold value of its spatial correlation with its nearby neighbors. Therefore some clusters may contain one cluster head, some 2, and some 5.

Another proposed functionality in the framework is the process of forming new clusters when unusual values are encountered. For example in the case of a forest fire, the network should be reinitialized and new clusters should be formed to better monitor the area. This is a necessary functionality because failing to Re-cluster can lead to incomplete data being sent, and could lead to highly damaged resources [5].

III.PROPOSED FRAMEWORK

The proposed framework extends functionality in sensor networks in which clusters are formed on the factor of spatial correlation. This framework is aimed at maximizing the lifetime of the sensor network. Specifically we take the example of the work done in [4] and add functionality to it.

A. Framework Overview

In this algorithm the spatial correlated weight between sensor nodes is taken which represent the degree of similarity between nearby sensor readings. Node that has the highest weight in its neighbours within communication radius becomes the Cluster Head. If nodes other than the Cluster Head have a correlated weight of more than 0.9, they are selected as the sub cluster heads. After all CHs have been selected, the non CH nodes join a cluster based on minimum distance. After the clusters are constructed, the data sampled in each cluster head have very high correlation with data sampled in each of its cluster members. Therefore only the CHs need to do the sampling work. This results in remarkable decrease in data transmitted without any extra data aggregation algorithm.

The cluster heads are selected in each round of sensing based on a transmit status and their remaining energy values. As environmental data is prone to high fluctuations, periodic re-clustering of the network takes place every 10 minutes. Re-clustering can also take place before the periodic 10 minutes if any abnormal value (outlier) is sensed.

B. Flowchart

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

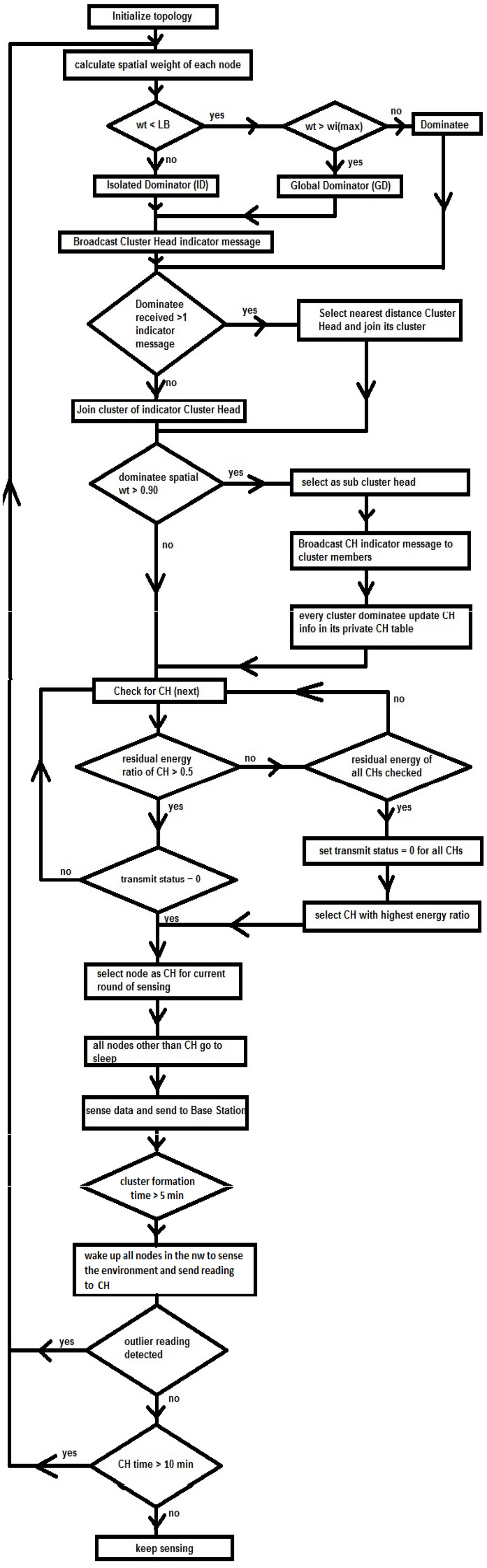

Fig. 1. Proposed Framework algorithm

The flowchart shown in Fig. 1 shows the flow of events in the proposed framework. The steps are as follows:

Step 1: Calculate the spatial correlated weight of each node.

Step 2: If wt < lower bound (assumed to be 0.70), make the node as isolated dominator.

Otherwise if nodes wt > all other nodes weight in its communication radius, make the node as a group dominator.

Step 3: Broadcast CH indicator message.

Step 4: All non Cluster Head nodes (dominatees) join a cluster. If the dominatee receives only one INDICATOR message, it joins the cluster of that cluster head. If it receives more than two INDICATOR messages, it joins the cluster of the nearest cluster head.

Step 5: Check for each dominatee node. If any dominatee has wt > 0.90, make it as a sub cluster head for that cluster.

Step 6: To select the cluster head for the current round of sensing, check if the residual energy of the CH is more than 0.50, otherwise check for the next. If it is 0.50, then check for its transmit status. If transmit status is 0, select it for the current round. Otherwise check for next cluster head. If all CHs have energy, 0.50 select node which has the highest energy ratio. If the transmit status of all CH is 1, reset their transmit status to 1.

Step 7: All non cluster head nodes go to sleep.

Step 8: Periodically every 5 minutes, wake up all the sensor nodes in the network to sense the readings one time. Every non cluster head node in the cluster sends its reading to its cluster head after sensing. If the cluster head node detects an outlier reading, call for the re-clustering algorithm at the base station otherwise put all non cluster heads again to sleep state.

Step 9: Check if CH time is greater than 10 min. If yes, activate the periodic re-clustering algorithm. Repeat Step 1.

C. Algorithms/Factors for proposed Framework

The algorithms/factors needed for this framework are:

- Spatial correlated weight
- Cluster head selection procedure
- Cluster construction procedure
- Sub cluster head selection procedure
- Periodic re-clustering time
- Outlier detection mechanism for re-clustering
- Residual energy ratio
- Transmit status for cluster head selection for each round of reading transmission

The different descriptions are described below.

i. Weight Calculation, Cluster Head Selection and Cluster Construction procedure

The weight calculation, cluster head selection procedure and cluster construction procedure have been defined by Ma et al. in [4]. This work is an extension to the above work, so the calculation of spatial correlated weight, cluster head selection procedure and construction of clusters from [4] has been used. To summarize what the authors have done:

The calculation of a node's weight given in equation (1) below tries to find out a measurement for each node to identify in what degree a node is correlated with other nodes in its communication radius i.e. if value of weight is high then the nodes are closeby and one node can represent its neighbors readings.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

d_{ij} represents the distance between two nodes.

Then the expected value of dij is $\overline{di}$
Where |N(i)| is the number of nodes in N(i).

$$w_i = \frac{[\sum_{j\in N(i)} |d_{ij} - \overline{d_i}|]^2}{|N(i)|^2 \; D(d_{ij})} = \frac{[\sum_{j\in N(i)} |d_{ij} - \overline{d_i}|]^2}{|N(i)| \sum_{j\in N(i)} (d_{ij} - \overline{d_i})^2} \qquad (1)$$

Ma et al. [4] have given the Cluster Head Selection procedure based on spatial correlated weight as follows:
A node decides to become a Cluster Head in two cases:

1. A node has very low correlation with all its neighbours.
2. A node has very high correlation with most of its neighbours.

The first kind of CH is known as a Isolated Dominator (ID) with no cluster members and the second type is known as a Group Dominator (GD) with some cluster members.
Step1: Calculate the spatial weight for each node
Step2: If (weight < lower bound), make it Isolated Dominator
Step3: Compare weights for all nodes along with their neighbors within communication radius. Node with the highest weight becomes the group dominator if there is no other group dominator in the group.
Step4: If a dominate node does not have a group dominator in its neighboring nodes, it becomes a group dominator.
After all the CHs are selected by the CHS procedure, each dominatee has to choose a cluster to join. In CC procedure, if a dominatee can be dominated by several dominators, it must choose the nearest dominator (the Euclidean distance is smallest between them) to join. If there are two cluster heads in the vicinity of the dominate node, it chooses the cluster of the nearest cluster head to join.
Please refer to [4] to understand the algorithms better.

ii. Sub-Cluster Head Selection

- After the cluster have been formed, for each cluster check for correlated weight of each member (dominatee)
- If node i is a dominatee with spatial correlated weight < spatial correlated weight of cluster head (dominator) but more than 0.9, set the dominatee as sub cluster head.
- There can be more than 1 sub cluster heads.

iii. Cluster Head Selection for Each Round of Sensing and Data Transmission

- Only one cluster head is active in each round of sensing.
- Alternate between cluster head and sub cluster heads
- Assign each cluster head transmit status a value of 0 at the beginning. As a cluster head transmits assign its transmit status to 1.
- At every reading transmission check for cluster head whose transmit status is 0 and residual energy is more than 0.5. if yes, let it transmit.
- If all cluster heads have transmit status to 1, reset transmit status of all to 0, and select the cluster head with most energy left to transmit.
- A cluster will not transmit if its residual energy is less than half of initial energy and other cluster heads have more than half of their initial energy.

iv. Periodic Re-clustering Mechanism

Periodic Re-clustering happens for better network sensing.

- Re-clustering time would be kept at 10 min for experimental purpose. This time chosen because phenomena such as forest fires happen occasionally. So keeping the time < 10 min would waste unnecessary energy of the sensing node. Also keeping the time as too long (e.g. 30 min or 1 hour) could result in forest fire to spread a long distance before being detected.

v. Outlier Re-clustering mechanism

- Periodically every 5 minutes, wake up all the sensor nodes in the network.
- Every node senses data and sends it to their respective cluster head.
- If the current cluster head or the sensing node detects abnormalities in the sensed reading than the usual trend, it calls for the re-clustering algorithm from the base station. Outlier detection technique described below is used to find out the abnormal readings. If not the non cluster head nodes again go to sleep.

vi. Abnormal Reading Detection for Re-clustering Mechanism

An event such as fire in the area would produce unusual sensed values. When an outlier reading is detected, the cluster construction phase is again restarted. Spatial correlated weights are again calculated for each and every node and new network structure may be formed. Every cluster head stores its last 10 readings from each non cluster node to calculate for future outlier readings.
An outlier detection method has been proposed in [4] depending on interquartile ranges (IQR). The IQR can be used as a measure of how spread-out the values are. It assumes that the values are clustered around some central value. Since in sensor networks environmental fluctuations are rare and generally a linear pattern of readings are detected, this method is appropriate for outlier detection in sensor networks.

1. Arrange the data points from lowest to highest. Calculate the median of the data set.
2. Calculate the lower quartile (Q1).
3. Calculate the upper quartile (Q3).
4. Find the "inner fences" for the data set. The first step is to multiply the difference between Q1 and Q3 (called the interquartile range) by 1.5. Add this number to Q3 and subtract it from Q1 to construct the inner fences.
5. Any data points that lie outside this range are considered outliers.

vii. Energy Factor

A cluster head can only be selected for current round of sensing and transmission based on energy parameter.

- E = En(current)/ En(initial)
- In the proposed protocol the ratio will be kept around 0.50 for experimental purpose.
- If multiple cluster heads are present only cluster heads whose energy is above 50% can transmit.
- Note: eventually if all cluster heads have energy levels below 0.50, then the cluster head with the maximum energy would transmit.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

IV. METRIC TO MEASURE LIFETIME

Time is used to measure the network lifetime. Two readings would be taken and analyzed for the algorithm efficiency.

- Time till 1st node dies (t1) : This will give a lower bound for the lifetime of the network
- Time till half nodes die (t[n/2])
- These metrics can be used to compare with different algorithm results to find out which one is more efficient.

V.EXPECTED OUTCOME

- The sensor network's lifetime with dynamic cluster heads is expected to be greater than the original network with a single cluster head.
- Re-clustering would help in better and accurate data being sent to the sink.
- Re-clustering based on abnormal values (outliers) could result in detection of unusual phenomena such as forest fires.

VI.CONCLUSIONS

In this paper a cluster formation algorithm with the use of dynamic cluster heads has been proposed. The proposed algorithm is expected to significantly improve the lifetime of the network, to sense the environmental data. Also a periodic re-clustering mechanism based on outlier's has been implemented which could help in tracking breakout of phenomena such as forest fires.

REFERENCES

[1] S. Lindsey and C.S. Raghavendra, "PEGASIS: Power Efficient Gathering in Sensor Information Systems," Proc. IEEE Aerospace Conf., pp. 1125-1130, 2002.

[2] S. Yoon and C. Shahabi. "Exploiting spatial correlation towards an energy efficient clustered aggregation technique (CAG)," IEEE International Conference on Communications, pages 3307-3313, May 2005.

[3] Chih-Chieh Hung, Wen-Chih Peng, Wang-Chien Lee, "Energy-Aware Set-Covering Approaches for Approximate Data Collection in Wireless Sensor Networks," IEEE Trans. Knowl. Data Eng, vol 24, no 11, 2012.

[4] Yajie Ma, Yike Guo, Xiangchuan Tian, and Moustafa Ghanem, "Distributed Clustering-Based Aggregation algorithm for Spatial correlated Sensor networks", IEEE Sensors Journal, Vol. 11, no. 3, pp 641-648, March 2011.

[5] Ashanie Guanathillake, Kithsiri Samarasinghe, "Energy Efficient Clustering Algorithm with Global & Local Re-clustering for Wireless Sensor Networks", International Journal of Communication Science and Engineering Vol. 7, No. 7, pp. 45-52, 2013.

[6] Wendi Rabiner Heinzelman, Anantha Chandrakasan, Hari Balakrishnan, "Energy Efficient Communication Protocol for Wireless Microsensor Networks", Proceedings of the 33rd Hawaii International Conference on System Sciences, IEEE, pp.1-10, 2000.

[7] Shakya, R.K. Singh, Y.N. ; Verma, N.K. "Generic correlation model for wireless sensor network applications" Wireless Sensor Systems, IET ,Vol 3 Issue: 4, pp. 266 – 276, ISSN: 2043-6386, Dec 2013

[8] Guangzhou, China; Yiju Zhan ; Yonghua Wang "Data Density Correlation Degree Clustering Method for Data Aggregation in WSN," Sensors Journal, IEEE, Vol 14 Issue, pp. 1089-1098, ISSN: 1530-437X, April 2014.

[9] Ghalib A., Shah, Muslim Bozyigit "Exploiting Energy aware Spatial Correlation in Wireless Sensor Networks," 2nd International Workshop on software for Sensor Networks (Sensor Wave 2007), 2007.

[10] http://en.wikipedia.org/wiki/Wireless_sensor_network accessed on 6.8.2013.

[11] C. Hua and T.-S. Yum, "Optimal routing and data aggregation for maximizing lifetime of wireless sensor networks," IEEE/ACM Trans. Netw. , vol. 16, no. 4, pp. 892–903, Aug. 2008

[12] D. Chu, A. Deshpande, J.M. Hellerstein, and W. Hong, "Approximate Data Collection in Sensor Networks Using Probabilistic Models," Proc. 22nd Int'l Conf. Data Eng. (ICDE), pp. 48-59, 2006

[13] O. Younis and S. Fahmy, "HEED: A Hybrid Energy-Efficient Distributed Clustering for Adhoc Sensor Networks," Proc. IEEE INFOCOM, 2004

[14] http://www.mathwords.com/o/outlier.htm accessed on 10.2.2014

Vivek Yadav is a student at SRM University Chennai. He is currently pursuing his M.Tech from SRM University, Chennai. His area of interests include Wireless and Mobile Communications.

Dr M. Pushpalatha is a professor at SRM University, Chennai. She completed her Ph.D in Computer Science and Engineering in 2014 from SRM University. Her area of interests include Wireless and Mobile Communications, Operating Systems and Artificial Intelligence and Expert Systems.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.