

AN ENERGY EFFICIENT DATA AGGREGATION FOR UNEQUAL CLUSTER WITH DOUBLE C-H BY POLYNOMIAL REGRESSION IN 3-D WIRELESS SENSOR NETWORK

DHARMESH YADAV¹, ANKIT TRIPATHI², Dr. ANSHUJ JAIN³

ELECTRONICS & COMMUNICATION,
SCOPE COLLEGE OF ENGINEERING, BHOPAL, MP, INDIA

Abstract

Sensor nodes observing events, occurring in close proximity, gather data that are highly co-related. This work uses a double cluster head and a polynomial regression technique to exploit the spatial correlation of the data in a three dimensional sensor network. The sensing nodes, sense the physical attribute and report their position coordinates (x, y, z) and the sensed value to the nearest tree node. Another set of nodes, categorized as the tree nodes, are responsible for generating a polynomial function of the received data and transmit the coefficients of regression to the parent tree node. The approach proceeds from the bottom to the top. The query from the sink, receives a polynomial function which is generated by the root node of the tree in each cluster to compute the attribute value at any location within the boundary. The proposed approach aims to save a lot of energy in the sensor network. Simulations, performed for different tree heights, indicate that a tree with a depth of four gives the best results.

Keywords: Data aggregation, Energy efficiency, Spatial correlation.

I. Introduction

Sensor nodes, in a wireless sensor network, are deployed to capture physical or environmental attribute. The wireless communication and some level of intelligence for signal processing and networking of the data, enables the sensor nodes to process the captured data and deliver it to the sink node in a multi hop fashion. The root node, also known as the sink, acts as an accumulator of all the data being sensed by the nodes. Sensor networks have different real life applications such as, habitat monitoring, military applications etc.

Large scale wireless sensor networks suffer from energy problems. Hence energy saving techniques are employed. The operation of wireless sensor networks depends on the data which is sensed and transmitted to the root node. The collection of sensor nodes monitor events and transmit their data. Dense deployment of sensor nodes allow some kind of correlation between the data being sensed by them. The two kinds of correlation that can exist in any given sensor network are:

- **Spatial Correlation:** Densely deployed sensor nodes produce data that are spatially correlated [9]. The distance between the nodes determines the degree of correlation.
- **Temporal Correlation:** Periodic observations from nodes generate data that shows temporal correlation. It has been observed that many physical phenomena show temporal correlation between each consecutive observation of a sensor node. The degree of correlation depends upon the temporal variation characteristics of the phenomenon.

In this paper, we have proposed a scheme that uses a polynomial regression technique to exploit the spatial correlation of the data in a three dimensional sensor network. The sensing nodes, sense the physical attribute and report their position coordinates (x, y, z) and the sensed value to the nearest tree node. Another set of nodes, categorized as the tree nodes, are responsible for generating a polynomial function of the received data and transmit the coefficients of regression to the parent tree node. The approach proceeds from the bottom to the top. The query from the sink, receives a polynomial function which is generated by the root node of the tree in each cluster to compute the attribute value at any location within the boundary.

The rest of the paper is organized under the following headings: section II describes the related work in the field of data aggregation in WSNs. The section III presents the proposed approach. The simulation results are presented in section IV. The section V conclude the proposed approach.

II. Previous Work

Data aggregation can be done via data aggregation tree (flat networks) or by a clustering strategy if the network is hierarchical. Considering the two types of networks, flat networks have equal nodes and connections are set up between nodes that are within each other's radio range, although constrained by connectivity conditions and available resources. A hierarchical network unlike the flat network has all nodes typically function both as switches/routers. One node in each cluster is selected as the cluster head (CH). In a hierarchical network, the number of tiers can vary according to the number of sensor nodes. The cluster head is responsible for routing the traffic, which essentially flows through the cluster head in every cluster. On the other hand, the gateway nodes are responsible for maintaining connectivity among neighboring cluster heads. The hierarchical network architecture is energy efficient for collecting and aggregating data from the entire WSN or all nodes within a larger target region, using knowledge of their relative locations, but the flat network architecture is often considered for transferring the sensor data between source-destination pairs separated by a large number of hops. In a large scale of network, multi-hop communication mode is usually adopted in order to save energy. In this scenario, the CHs closer to the BS will take on more forwarding tasks, which results in heavy overhead of CHs, and these CHs will run out of power earlier than other nodes. It causes breakdown of the cluster and the communication lost between CHs. This is called "hotspots" problem

Data correlation is an extensive field of research and a huge amount of literature exists to study data correlation in WSN. The authors in [3], [4], [5] explore the theoretical aspects of the correlation. The aim of the study of these works is to find the optimum rate to compress redundant information in the sensor observations. A recent work in [6], studies the relation between spatiotemporal bandwidth, distortion, and power in large scale WSNs.

Recent years have witnessed extensive research on in-network aggregation for WSNs [5, 8, and 9]. Enachescu et al. in [10] propose a very simple opportunistic aggregation scheme with near optimum performance under widely varying scales of correlation. The scheme, proposed as Opportunistic Aggregation scheme, that sends the data from a sensor to the central agent over a shortest path. The authors formalize a notion of correlation that can vary according to a parameter k . The expected collision time of "nearby" walks on the grid to the optimum cost of scale-free aggregation is then related. A randomized algorithm is used for routing information on a grid of sensors that satisfies the appropriate collision time condition. The proposed approach thus proves that the proposed scheme is a constant factor approximation (in expectation) to the optimum aggregation tree simultaneously for all correlation parameters k . In-line data aggregation is essential for WSNs where the energy resources are limited. Intanagonwiwat et al. have proposed an aggregation scheme in [10]. The authors propose an approach based on adjusting the aggregation points which results in increased path sharing and reduced overall consumption of energy. First a greedy incremental tree is created which is responsible for the establishment of a shortest path to the first source to the sink and hence the name, greedy. The other sources are incrementally connected at the closest point on the existing tree. The results show that greedy aggregation can achieve up to 45% energy savings over opportunistic aggregation in high-density networks without adversely impacting latency or robustness.

Zhong et al. have presented an ultra-low power MAC scheme. When not in use, the radio devices can be turned off to save energy. The advantage of this approach is its ability to take advantage of the redundant data built in the network. The lost data can be efficiently recovered at the destination on the basis of the data from the lost data's neighborhood. The high spatial correlation between the data in close proximity allows the recovery from the vast data available.

A Tree based polynomial regression algorithm, (TREG) is proposed by Banerjee et al. in [7]. The method proposed in [7] is based on the degree of correlation between the sensor data. The authors propose to use a multiple attribute-based binary tree, which with the help of a polynomial regression function, generates aggregated data at the root node. The proposed approach is limited to 2D space in which the nodes sensing the data, run a regression function to find a polynomial in the space coordinates (x, y) . The coefficients of the function are forwarded by each node and hence the raw data need not be sent. At every level, a regression polynomial is found increasing the efficiency of the approximation function. The transmission of coefficients, instead of the raw data, allows a constant data packet size for all the nodes, thereby saving a lot of energy.

In this paper, a novel data aggregation algorithm has been proposed to exploit the spatial correlation of the data in a three dimensional sensor network by using polynomial regression technique. The proposed algorithm is analysed on the basis of simulation results which prove the proposed algorithm to be better and efficient than the existing approaches for data aggregation in a 3D space.

III. Proposed Work

The proposed scheme uses an energy-efficient routing protocol which adopts unequal clustering technology to solve the hot spots problem and proposes double cluster head strategy to reduce the energy

consumption of head nodes in the clusters. In this scheme, each cluster has two CHs. Main head is responsible for forwarding data between the clusters and vice head is responsible for receiving and aggregating data within each cluster. For this polynomial function based regression technique is used which results effective and energy efficient data aggregation in WSN. In addition, to balance the energy consumption between cluster heads and cluster members, a hybrid cluster head rotation strategy based on time- driven and energy-driven is proposed. The technique is effective for three dimensional space, and hence a wide area of application. In the present work, the sensing field is divided into a number of clusters. Each cluster constructs a tree within the cluster, through which the data is propagated to the root node. The simulation results for randomly deployed sensor nodes are presented with different node densities.

A. Assumptions

We make the following assumptions in our work

1. The sensor nodes are stationary.
2. Each node knows its location coordinates (x, y, z) in the three dimensional space.
3. For all the simulations, we have assumed a binary tree.
4. The MAC protocol is collision free.
5. Energy consumed for computation is negligible as compared to energy consumed for data transmission and data reception [17].
6. The base station not limits to energy.
7. Sensor nodes know the relative position of the base station in the field.
8. The sensor nodes are randomly distributed over the sensing field.
9. The Sensor nodes are densely deployed in the sensing field. This dense deployment of sensor network achieving Quality of Service.
10. Sensor nodes are homogeneous in their architectures.
11. Sensor nodes are able to measure the current energy level.

B. Description of the scheme:

Our scheme is based on the fact that there is a correlation between attribute values and the location. Multiple sensors, in close proximity, can detect the same event and give almost similar readings. The basic objective here, being the aggregation of the sensed data in such a way so as to reduce the data redundancy.

a) Clustering

The clustering includes cluster head selection and cluster formation process.

In the **cluster head selection process** each sensor node chooses a random number between 0 and 1 separately. A node becomes the cluster head, if this number is lower than the calculated threshold for the node.

In the **cluster formation process** each cluster head broadcasts a join message within the sensing field. On reception of the "join" message each non cluster head sensor node, decides to join the cluster head, if it receives the join message more than once then the sensor node joins the nearest cluster head. After a constant time interval cluster head receives join request messages from all the non-cluster head sensor nodes which intend to join it.

After cluster formation each cluster head work as the root node and create a tree in cluster. Cluster head randomly choose the sensor nodes for tree formation.

Vice CH selection. In order to alleviate the task load of CHs, we propose a double-CH mechanism. The main CH is responsible for communication, and the vice CH is responsible for data collection and maintenance within the cluster. The concept of VCH in the literature is presented to extend the network lifetime and handle the nodes in a cluster after the inability of the current CHs to perform. Mostly, the node having second highest residual energy in a cluster is selected as the VCH and it replaces CH's functions in case of its energy drain. The node with the highest rank in a cluster is selected as the CH. After selection, the CH announces itself in the cluster by broadcasting a message. The member nodes send ACK messages along with their ranks. The CH chooses the node with the highest rank as the VCH and declares it. After clusters set up phase, each vice CH creates a TDMA schedule and broadcasts schedule to its all CMs. Based on the allocated time slot, each CM collects and sends data to its Parent node and then node send this data to its corresponding vice CH. In order to reduce the communication burden, the vice CH applies aggregation technique to process the data once it receives data from all Parent nodes. Then the aggregated data are sent to the corresponding main CH. Finally, main CHs transmit the data to the BS by multi-hop routing.

b) Tree Formation Algorithm

A tree is formed within the cluster through which the sensed data is propagated to the root node. The steps of the tree formation algorithm can be summarized as:

1. Select the Vice cluster head node as as the root of the binary tree. Using the value of tree height (h), the number of nodes in the tree is computed.
2. $NumberofTreeNodes = Nt = 2(h + 1) - 1$
3. The number of tree nodes to be selected = $(Nt - 1)$.
4. Randomly select $(Nt - 1)$ sensor nodes.
5. The distance for the randomly-selected node is then computed (from the root node).
6. Two nodes with minimum distance from the root node (closest to the root node) will become its children.
7. For all the remaining nodes compute their distance from both the newly assigned child nodes.
8. Two nodes closest to the first node will become its children.
9. Similarly, two nodes closest to the second node, will become its children.
10. The process terminates when the tree formation is complete.

The tree formation takes place from top to bottom. It starts with the root node that we initialize first. The number of tree nodes is randomly selected. But, how and at what level of the binary tree a node will join the tree is determined by finding the distance between the nodes. Initially the root has two child positions available. Out of the selected nodes, two nearest to the root will become its children. Then from the remaining nodes, the two nearest to each child will become that child's children. The algorithm thus progresses from top to bottom till the tree formation is complete.

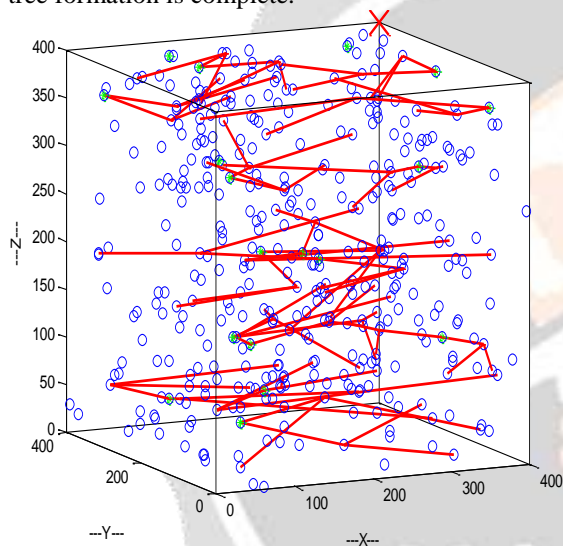


Figure 1: Tree formation in cluster (Height 2)

After completion the tree formation we have three type of nodes in the network viz. the root node, the tree nodes and the sensing nodes

The sensing nodes are responsible for capturing the attribute value and report it to the nearest tree node. Each data packet, sent to the tree node, consists of:

- Sensed value of the parameter
- Its coordinates in space (x, y, z)

Many nodes send their data to the tree node. The tree node, on receiving the data from multiple sensing nodes, runs the polynomial regression function to get the coefficients of regression. The polynomial function in our approach is given by:

$$t = F(x, y, z) \quad (1)$$

$$F(x, y, z) = b_0 + b_1 z + b_2 y + b_3 yz + b_4 x + b_5 xz + b_6 xy + b_7 xyz(2)$$

Each tree node, then computes the range of coordinates (using the values of the coordinates sent from different sensing nodes). The computed range is given by $(x_{\min}, x_{\max}, y_{\min}, y_{\max}, z_{\min}, z_{\max})$. Thus the minimum and maximum value of x, y and z components of the coordinates is calculated. The eight coefficients of regression, are then sent to the parent node, as well as the range of coordinates for which the coefficients have been computed. The parent tree node on receiving the coefficients from its children has the following three sets of data

1. Coefficients of regression + Range of Coordinates sent from first child

2. Coefficients of regression + Range of Coordinates sent from second child
3. Sensed data + Coordinates sent from nearby sensing nodes.

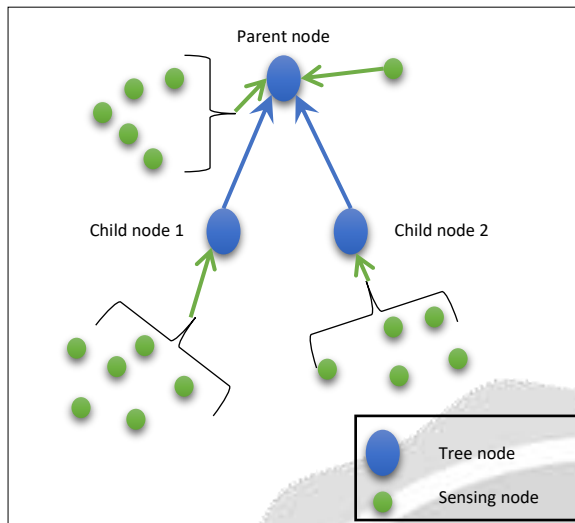
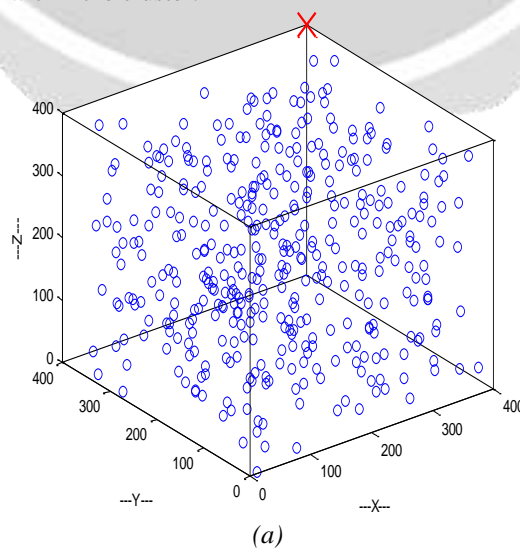


Figure 2: Data transmission between nodes

The parent tree node uses the data set 1 to randomly generate points (x, y, z) such that each generated point lies within the range of coordinates received in data set 1. It then uses the coefficients of regression received in data set 1 to compute the attribute values at the randomly generated points. So ultimately it gets the data (x, y, z, t) where “t” is the attribute value at point (x, y, z) in space. It does a similar computation for the second set of data. The third set of data that it received from the sensing nodes, is already in the form (x, y, z, t) . Now the parent node uses the combined data set for polynomial regression. It computes the coefficients of regression for the same polynomial function and the combined set of data. It also computes the overall range of coordinates from the three sets of data. It then sends the coefficients of regression and the range of coordinates to its parent tree node. This process starts at the leaf nodes and stops at the root (Vice cluster head). At the end of the process the root has the polynomial function that represents the attribute value as a function of coordinates in space. Instead of querying the network for attribute value at any point (x, y, z) , the root node can now use this polynomial function to compute the attribute value for any node in the network using its (x, y, z) value.

IV. Simulation Results

The simulation results for randomly deployed sensor nodes in a regular grid pattern in a $400\text{ m} \times 400\text{ m} \times 400\text{ m}$ volume of space. The figure 3 (a) shows a sensor field in which the sensors have been deployed randomly and the figure 3 (b) represents the cluster head selected in sensing field. The cluster head act as the root node for every tree formed within the cluster.



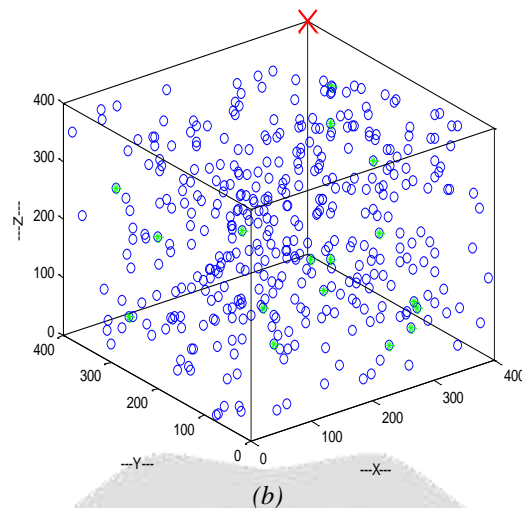


Figure 3 (a) Randomly Deployed Sensornodes. (b) cluster head as root node in cluster

A randomly generated temperature value which is spatially correlated in a region is used for the purpose of simulations.

The following parameters were assumed as:

Table 1: Simulation parameters

<i>Parameter</i>	<i>Value</i>
The energy consumed to receive 1 bit	10 pJ
Initial energy of each sensor node	1 mJ
The energy consumed to transmit 1 bit	10 pJ

A. Results for randomly deployed network

Simulations are performed for different treeheights, from 2 to 5 in a varying network density (i.e. number of nodes in the network). The results are evaluated under the following metrics:

1. The accuracy of the approximation of the sensed parameter over the entire region in percentage error
2. Percentage of Compression achieved
3. Energy Consumed in the tree nodes.

a) Percentage Error.

We computed the percentage error as follows

$$Error = \frac{|T_{actual} - T_{computed}|}{T_{actual}} \times 100(3)$$

Where T_{actual} is the actual temperature value at a point and $T_{computed}$ is the computed temperature with proposed scheme.

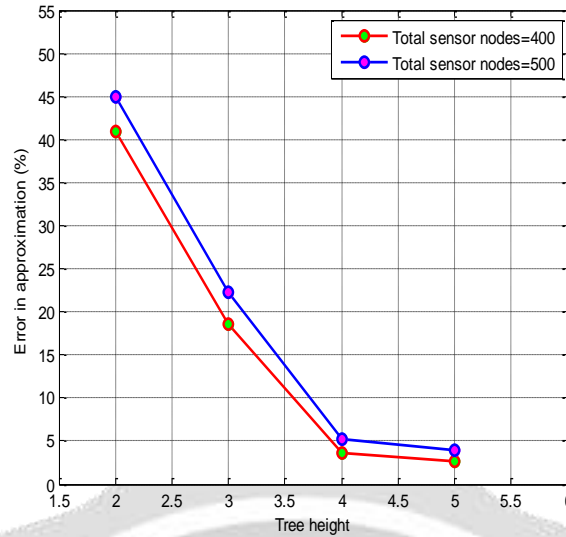


Figure 4. Percentage Error vs. TreeHeight

In Figure 4, shows a decrease in the percentage error as the height of the tree increases. As the tree height increases, a larger area can be covered leading to a better sensed parameter over the region. The maximum error is obtained with the tree with depth 2. A tree with a height of 4, covers the entire network and is able to approximate the network more efficiently. Although there is not much significant improvement in accuracy as the tree height increases from 4 to 5.

b) Data Compression.

Data packets of the same size, is transmitted by each node to their parent node. This transmission is independent of the number of sensing nodes reporting their data to the tree node. The regression prohibits every tree node to transmit all the data that it received from sensing nodes. Percentage compression is computed and is given by:

$$Compression = \frac{|B_{input} - B_{output}|}{B_{input}} \times 100 \quad (4)$$

Where, B_{output} and B_{input} are the bits transmitted and bits received by the tree nodes respectively. It is observed that compression increases with an increase in the number of nodes in the network increase.

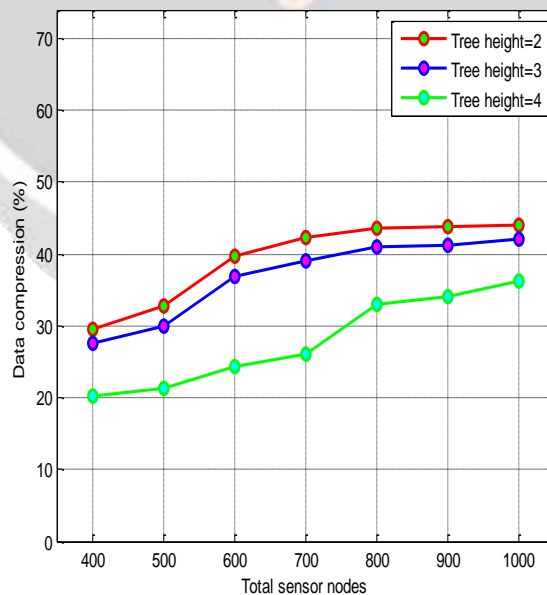


Figure 5. Percentage Compression vs. Number of Nodes

c) Energy Consumed in Tree Nodes.

Calculation of energy consumed in tree nodes is based on the assumption that all the tree nodes have already received the data from the sensing nodes. The graph indicates that the energy consumed in the tree is not

dependent on the number of nodes in the network. This is expected as number of transmissions in the tree remains the same for a given tree height.

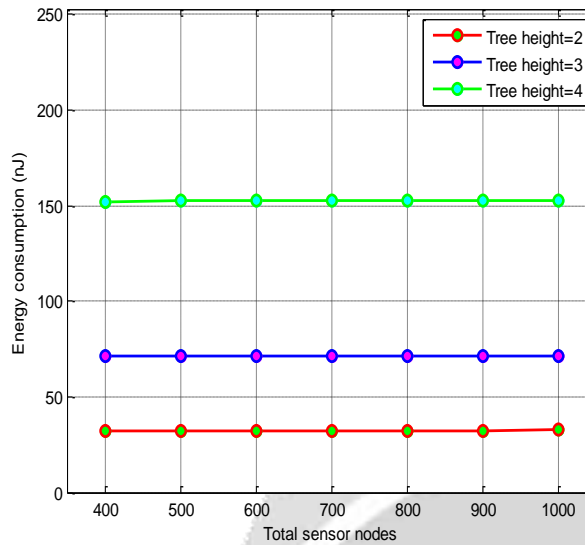
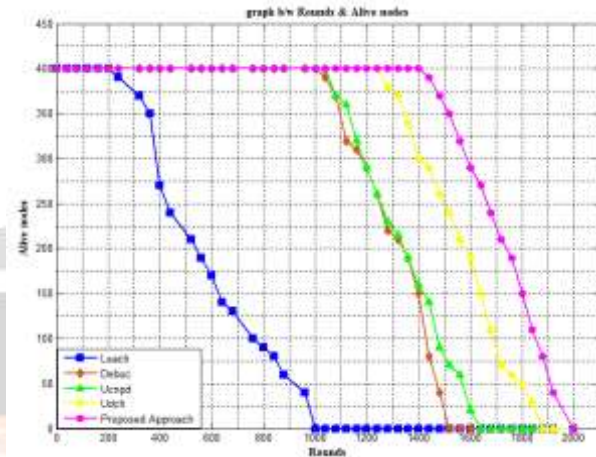


Figure 6. Energy Consumed in TreeNodes



d) Comparison with two-dimensional polynomial regression scheme.

The Figure 7, compares, a two dimensional TREG scheme [7] with the proposed scheme. Results prove that it is more in percentage error as the two dimensional TREG scheme considers only the two dimensions in the sensing space.

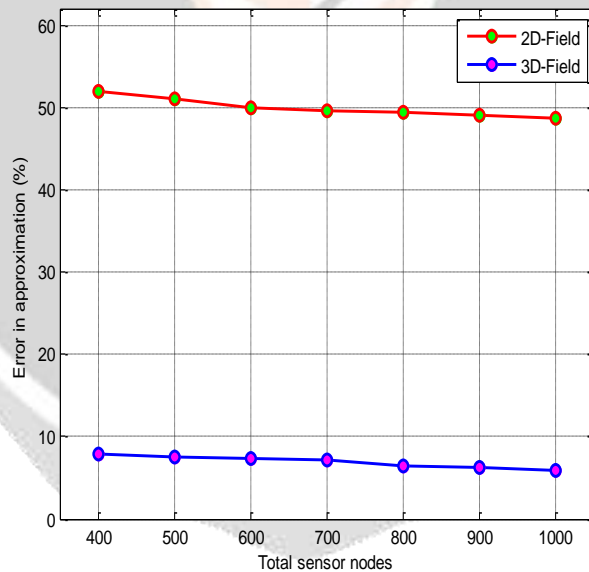


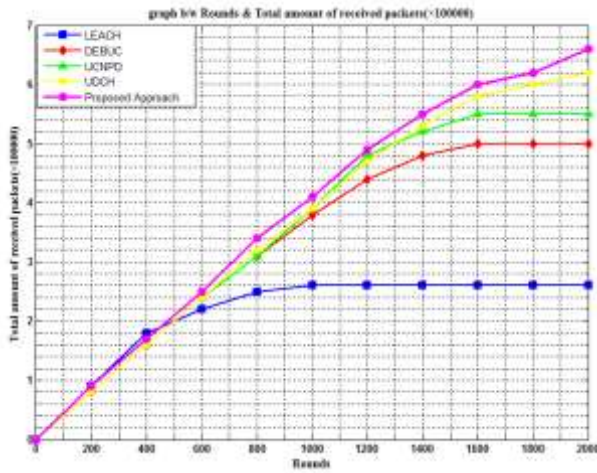
Figure 7. Two Dimensional Approximation versus three Dimensional

e) Lifetime comparison.

Network lifetime is very important indicator for evaluating energy efficiency performance. The time of the first death node appearance also can reflect stability of network to a certain extent. In this work, we take the death of all nodes as end of network life- time. The network lifetime of four protocols is compared as shown in Figure: 5.10 The simulation results show that lifetime of LEACH is about 1050 rounds, DEBUC is about 1550 rounds, UCNPD is about 1630 rounds, and UDCH is about 1870 rounds and for proposed Approach is about 1990. The lifetime of LEACH is shortest due to the redundant data and unbalanced energy consumption. In proposed Approach double-CH mechanism with polynomial function based regression technique is used therefore, redundant data and energy balance are improved to some extent, and their lifetime is longer than other protocols.

f) Throughput comparison

Throughput is defined as the number of data packets successful received at BS. Figure shows results of simulation where the throughput of given protocols is compared. The results show that LEACH receives about 260,000 packets totally during 2000 rounds, DEBUC receives about 500,000 packets, UCNPD receives about

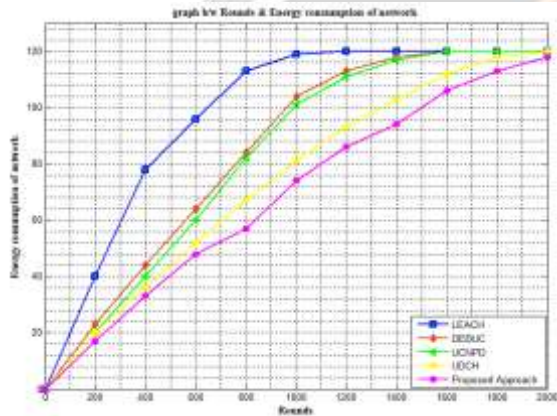


550,000 packets, UDCH receives about 610,000 packets and Proposed Approach receives about 660,000 packets. LEACH receives most packets before first 500 rounds due to large number of redundant data. After 500 rounds, dead nodes appear in LEACH; therefore, data packet begins to reduce. DEBUC and UCNPD adopt unequal cluster technology, so the distribution of the cluster is more reasonable and energy consumption is more balanced; the total amount of data packets is more than LEACH. The death of sensor nodes is slowest in Proposed Approach and the lifetime is the longest; therefore, the total amount of packets is largest.

g) Energy consumption comparison

Energy consumption implicitly reflects the status of network lifetime. The simulation results of energy consumption are shown in Figure 5.12. The energy consumption of LEACH is larger than the other three protocols due to random cluster distribution and redundant packets' transmissions. DEBUC and UCNPD adopt CH selection algorithm

Based on residual energy, which leads to more reasonable cluster distribution and more balanced energy consumption; therefore, during the same time, energy consumption is smaller than LEACH. UDCH improves CH selection algorithm and competition radius calculation algorithm, which makes the distribution of clusters more reasonable.



V. Performance Analysis

A. Advantages

The scheme has following advantages:

1. For a polynomial with eight coefficients, the polynomial regression function requires eight unique and independent sets of data. If the size of the data set increases, unlike the previous approaches, the proposed approach partitions the data and generates a set of coefficients, which are individually sent. Even if the size of the data set decreases, the proposed approach can efficiently generate the coefficients.
2. Reduced number of transmissions and therefore reduced energy consumption.
3. The coefficients of regression and minimum and maximum coordinates are only sent to the parent node. Reduced data size saves energy.
4. The regression scheme when employed on the data and reduces the redundancy.

B. Limitations

The limitations of the proposed approach are:

1. Results are presented for the specific temperature model. A different set may produce different results
2. A different set might require different function to efficiently approximate the pattern.

VI. Conclusions and Future Work

The Proposed scheme exploits the spatial correlation of attributes in a three-dimensional sensor network. The heart of the proposed scheme is an attribute-based tree in the clustered network and polynomial regression at every tree node when a dimension of the data is less. The root node computes the value of physical attribute at any point in the sensor network by inputting the coefficients of that point in a polynomial function. The simulation result shows that the proposed technique performs better than the previous technique.

References

1. [1] D. P. Agrawal and Q. A. Zeng, *Introduction to Wireless and Mobile Systems*, 436 pages, Brooks/Cole Publication, 2003.
2. [2] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, 43(5):51–58, May 2000.
3. [3] M. C. Vuran, O. B. Akan, and I. F. Akyildiz, "Spatio- Temporal Correlation: Theory and Applications Wireless Sensor Networks," *Computer Networks Journal (Elsevier Science)*, vol. 45, no. 3, pp. 245 - 259, June 2004.
4. [4] C. Intanagonwiwat, D. Estrin, R. Govindan and J. Heidemann, "Impact of Network Density on Data Aggregation in Wireless Sensor Networks," University of Southern California, Information Sciences Institute.
5. [5] L. Zhong, R. Shah, C. Guo, and J. Rabaey, "An ultralowpower and distributed access protocol for broadband wireless sensor networks," presented at the Networkworld + Interop: IEEE Broadband Wireless Summit, Las Vegas, NV, May 2001.
6. [6] H. S. Kim and W. H. Kwon, "Spatial and Temporal Multi-Aggregation for State-Based Sensor Data in Wireless Sensor Networks," *Telecommunication Systems, Springer Netherlands Volume 26, Numbers 2-4 / June, 2004* 161-179.
7. [7] T. Banerjee, K. Choudhury and D. P. Agrawal, "Tree
8. Based Data Aggregation in Sensor Networks Using Polynomial Regression," *Proceedings of the 8th Ann. Conf. on Information Fusion, July 25 - 29, 2005*,
9. [10] J. Hill, R. Szewczyk, A. Woo, S. Hollar, D. E. Culler, and K. S. J. Pister, "System architecture directions for networked sensors," In *Proc. of the 9th International Conference on Architectural Support for Programming Languages and Operating Systems*, page 93-104, Boston, MA, USA, Nov. 2000.
10. Zijng Wang, Xiaoqi Qin, Baoling Liu. "An Energy-Efficient Clustering Routing Algorithm for WSN-Assisted IoT" *IEEE Wireless Communications and Networking Conference(WCNC)*,2018
11. H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. H. Hanzo, "A survey of network lifetime maximization techniques in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 828-854, Second Quarter, 2017
12. Karjee, Jyotirmoy, Martin Kleasinsteuber, and R. S. Jamadagni. "Spatial data estimation in three dimensional distributed wireless sensor networks." *Embedded Systems (ICES), 2014 International Conference on. IEEE, 2014.*
13. Karjee, Jyotirmoy, and H. S. Jamadagni. "Optimal Node Selection Using Estimated Data Accuracy Model in Wireless Sensor Networks." *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing. Springer New York, 2013.*
14. S.Prasanna and Srinivasa Rao "An Overview of Wireless Sensor Networks Applications and Security" *International Journal of Soft Computing and Engineering (IJSCE) Volume-2, Issue-2, May 2012.*
15. Ranjani, S. Siva, S. Radha Krishnan, and C. Thangaraj. "Energy-efficient cluster based data aggregation for wireless sensor networks." *Recent Advances in Computing and Software Systems (RACSS), 2012 International Conference on. IEEE, 2012.*
16. deMoraisCordeiro, Carlos, and Dharma Prakash Agrawal. *Ad hoc and sensor networks: theory and applications.* World scientific, 2011.
17. Yigitel, M. Aykut, OzlemDurmazIncel, and CemErsoy. "QoS-aware MAC protocols for wireless sensor networks: A survey." *Computer Networks* 55.8 (2011): 1982-2004.
18. Choi, Jinchul, and Chaewoo Lee. "Energy consumption and lifetime analysis in clustered multi-hop wireless sensor networks using the probabilistic cluster-head selection method." *EURASIP Journal on Wireless Communications and Networking* 2011.1 (2011): 1-13.
19. Ozdemir, Suat, and Yang Xiao. "Polynomial regression based secure data aggregation for wireless sensor networks." *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE. IEEE, 2011.*
20. Patil, Nandini S., and P. R. Patil. "Data aggregation in wireless sensor network." *Conference on Computational Intelligence and Computing Research.* 2010.
21. Chen, Zhikui, et al. "A clustering approximation mechanism based on data spatial correlation in wireless sensor networks." *Wireless Telecommunications Symposium (WTS), 2010. IEEE, 2010.*
22. Dabirmoghaddam, Ali, Majid Ghaderi, and Carey Williamson. "Energy-efficient clustering in wireless sensor networks with spatially correlated data." *IEEE infocom.* 2010.

23. Ozdemir, Suat, and Yang Xiao. "Secure data aggregation in wireless sensor networks: A comprehensive overview." *Computer Networks* 53.12 (2009): 2022-2037.
24. Cai, Kang, Gang Wei, and Huifang Li. "Information accuracy versus jointly sensing nodes in wireless sensor networks." *Circuits and Systems*, 2008. APCCAS 2008. IEEE Asia Pacific Conference on. IEEE, 2008.
25. Bhattacharjee S, Das N. Distributed data gathering scheduling in multihop wireless sensor networks for improved lifetime. In: *International Conference on Computing: Theory and Applications (ICCTA)*, 2007.
26. Abbasi, Ameer Ahmed, and Mohamed Younis. "A survey on clustering algorithms for wireless sensor networks." *Computer communications* 30.14 (2007): 2826-2841.
27. Qing, Li, Qingxin Zhu, and Mingwen Wang. "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks." *Computer communications* 29.12 (2006): 2230-2237.
28. T. Banerjee, K. Choudhury and D. P. Agrawal, "Using Polynomial Regression for Data Representation in Wireless Sensor Networks," in *Special Issue of Wiley International Journal of Communication Systems on Energy-Efficient Network Protocols and Algorithms for Wireless Sensor Networks*, 2006.
29. M. Enachescu, A. Goel, R. Govindan and R. Motwani, "Scale Free Aggregation in Sensor Networks," Preprint submitted to Elsevier Science 24 January 2006.
30. Chen G, Guo T-D, Yang W-G, Zhao T. An improved ant-based routing protocol in wireless sensor networks. In: *International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, 2006.

