Enhance the Efficiency of Wireless Data Aggregation by Optimizing the Structure through EM Algorithm

Harpriya Guru[1], NavneetKaur[2]
[1]*M.Tech/System Software, Chandigarh Engineering College, Landran, Punjab Technical University*
[2]*Professor/CSE, Chandigarh Engineering College, Landran, Punjab Technical University*

Abstract— In wireless sensor use for aggregate and route the information node by node, but routing reduce the energy step by step ,so reduce the energy of nodes and increase the un stability of network, that's why basic challenge in WSN is reducing the energy loss at time of routing and data aggregation. In this paper reducing energy loss and increasing throughput by Expectation maximization algorithm. Which optimize the data aggregation?

Keywords— *WSN, optimization, EM, Data aggregation.*

I. INTRODUCTION

Recent advances in wireless communications leads to Wireless Sensor Network WSN. Basically WSNs possess some characteristics which include node deployment, limited battery power and memory, single or multiple base Station (BS), node dynamicity, no global unique ID, application awareness, immovable sensor nodes. The WSNs amazing features direct us to innovative research challenges in several data mining process.

In recent years, wireless sensor networks (WSNs) have emerged as a new category of networking systems with limited computing, communication, and storage resources. A WSN consists of nodes deployed to sense physical or environmental conditions for a wide range of applications, such as environment monitoring scientific observation emergency detection field surveillance and structure monitoring. Wireless sensor networks (WSNs) are increasingly used in many applications, such as volcano and fire monitoring, urban sensing, and perimeter surveillance. In a large WSN, in-network data aggregation (i.e., combining partial results at intermediate nodes during message routing) significantly reduces the amount of communication overhead and energy consumption.

In most cases, the sensor nodes form a multi-hop network while the base station (BS) acts as the central point of control. Typically, a sensor node has limitation in terms of computation capability and energy. One common way is to allow each sensor node to forward its reading to the BS, possibly via other intermediate nodes. In-network data aggregation can reduce the amount of communication and hence the energy consumed, especially in large WSNs. The main idea is to combine partial results at intermediate nodes during message routing. One approach is to construct a spanning tree rooted at the BS and then perform in-network aggregation along the tree.

Data aggregation in wireless sensor network is very much attracted by research communities nowadays to prolong the network lifetime. Data aggregation algorithms are frequently measured by executing the algorithm several rounds. Data aggregation algorithms are classified based on the type of communication architecture that it employs such as clusters, grid, chain, connected dominating sets and trees.

Data aggregation has been put forward as an essential paradigm for wireless routing in sensor networks. The idea is to combine the data coming from different sources enroute – eliminating redundancy, minimizing the number of transmissions and thus saving energy. This paradigm shifts the focus from the traditional address centric approaches for networking (finding short routes between pairs of addressable end-nodes) to a more data centric approach (finding routes from multiple sources to a single destination that allows in-network consolidation of redundant data).

Types of Aggregation Techniques

There are many types of aggregation techniques are present some of them are listed below.

1. Centralized Approach: This is an address centric approach where each node sends data to a central node via the shortest possible route using a multihopwireless protocol. The sensor nodes simply send the data packets to a leader, which is the powerful node. The leader aggregates the data which can be queried. Each intermediate node has to send the data packets addressed to leader from the child nodes. So a large number of messages have to be transmitted for a query in the best case equal to the sum of external path lengths for each node.
2. In-Network Aggregation: In-network aggregation is the global process of gathering and routing information through a multi-hop network, processing data at intermediate nodes with the objective of reducing resource consumption (in particular energy), thereby increasing network lifetime. There are two approaches for in-network aggregation: with size reduction and without size reduction. In-network aggregation with size

reduction refers to the process of combining & compressing the data packets received by a node from its neighbors in order to reduce the packet length to be transmitted or forwarded towards sink.

3. Tree-Based Approach: In the tree-based approach perform aggregation by constructing an aggregation tree, which could be a minimum spanning tree, rooted at sink and source nodes are considered as leaves. Each node has a parent node to forward its data. Flow of data starts from leaves nodes up to the sink and therein the aggregation done by parent nodes.
4. Cluster-Based Approach: In cluster-based approach, whole network is divided in to several clusters. Each cluster has a cluster-head which is selected among cluster members. Clusterheads do the role of aggregator which aggregate data received from cluster members locally and then transmit the result to sink.

Advantages of Data Aggregation in Wireless Sensor Network

With the help of data aggregation process we can enhance the robustness and accuracy of information which is obtained by entire network, certain redundancy exists in the data collected from sensor nodes thus data fusion processing is needed to reduce the redundant information. Another advantage is those reduces the traffic load and conserve energy of the sensors.

Disadvantages of Data Aggregation in Wireless Sensor Network

The cluster head means data aggregator nodes send fuse these data to the base station .this cluster head or aggregator node may be attacked by malicious attacker. If a cluster head is compromised, then the base station (sink) cannot be ensure the correctness of the aggregate data that has been send to it. Another drawback is existing systems are several copies of the aggregate result may be sent to the base station (sink) by uncompromised nodes.It increase the power consumed at these nodes.

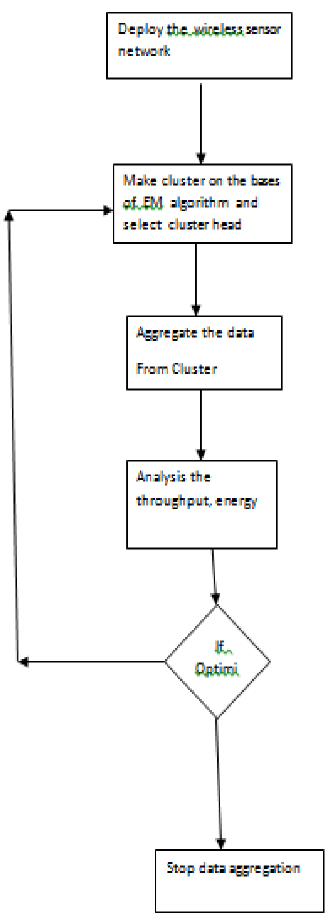
II. LITERATURE REVIEW

In [1] Ozyurt et al: In this paper the author proposed the nearest insertion method, where the farthest node is chosen first to be connected with a route. Then, repeatedly, each selected node chooses the nearest neighbor that has not been assigned a route so far, and connects itself to this neighbor. This procedure repeats until all customers are connected by routes.In [2] Solomon: In this paper the author developed the push forward insertion heuristic (PFIH), which repeatedly selects the customer with the lowest additional insertion cost as the next node, until all customers are connected. Once initial routes have been found, various algorithms are developed to generate near optimal solutions based on simulated annealing tabu search or genetic programming. In [3] Yanjun Yao et al: In this paper stems from the insight that, recent research efforts on open vehicle routing (OVR) problems, an active area in operations research, are based on similar assumptions and constraints compared to sensor networks. This insight motivates us to adapt these techniques so that they can solve or prove certain challenging problems in WSN applications. To demonstrate that this approach is feasible, they develop one data collection protocol called EDAL, which stands for Energy-efficient Delay aware Lifetime-balancing data collection. The algorithm design of EDAL borrows one research result from OVR to prove that its problem formulation is inherently NP-hard. they then proposed both a centralized heuristic to reduce its computational overhead, and a distributed heuristic to make the algorithm scalable for large scale network operations. They also develop EDAL to be closely integrated with compressive sensing, an emerging technique that promises considerable reduction in total traffic cost for collecting sensor readings under loose delay bounds. Finally, they systematically evaluate EDAL to demonstrate its performance superiority compared to related protocols.

In [4] Sankardas Roy et al: In this paper, they make the synopsis diffusion approach secure against the above attack launched by compromised nodes. In particular, they present an algorithm to enable the base station to securely compute predicate count or sum even in the presence of such an attack. Their attack-resilient computation algorithm computes the true aggregate by filtering out the contributions of compromised nodes in the aggregation hierarchy. Extensive analysis and simulation study show that their algorithm outperforms other existing approaches. In [5] Thorsten Nowak et al: In this paper a signaling scheme for a combined localization and communication using a common set of subcarriers is proposed. The concept is based on binary offset carrier signals. But, in contrast to Global Navigation Satellite Systems the presented approach makes use of pure subcarrier localization, and thus enables data transmission in short burst signals. The ranging performance is assessed utilizing the Cramer-RaoLower ´ Bound depending on the amount of data transferred and considering bit errors.

In [6] BhaskarKrishnamachari et al: In this paper they model data-centric routing and compare its performance with traditional end-to-end routing schemes. They examine the impact of source-destination placement and communication network density on the energy costs and delay associated with data aggregation. They show that data-centric routing offers significant performance gains across a wide range of operational scenarios. They also examine the complexity of optimal data aggregation, showing that although it is an NPhard problem in general, there exist useful polynomial-time special cases. In [7] Rajagopalan et al: In this paper they focus on data aggregation problems in energy constrained sensor networks. The main goal of data aggregation algorithms is to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced. They also present a survey of data aggregation algorithms in wireless sensor networks. They compare and contrast different algorithms on the basis of performance measures such as lifetime, latency and data accuracy. Finally, they conclude with possible future research directions.In [8] Nandini. S. Patil et al: In this paper, a data aggregation framework on wireless sensor networks is presented. The framework works as a middleware for aggregating data measured by a number of nodes within a network. The aim of the proposed work is to compare the performance of TAG in terms of energy efficiency in comparison with and without data aggregation in wireless sensor networks and to assess the suitability. Wireless sensor networks (WSN) offer an increasingly Sensor nodes need less power for processing as compared to transmitting data. It is preferable to do in network processing inside network and reduce packet size. One such approach is data aggregation which attractive method of data gathering in distributed system architectures and dynamic access via wireless connectivity.

III. METHODOLOGY



Step1: Deploy the wireless sensor network.
Step2: Make cluster on the bases of EM algorithm and select cluster head.
Step3: Aggregate the data from cluster.
Step4: Analysis the throughput and energy.
Step5: In this process optimization is done.
Step6: If data is optimized then stop data aggregation otherwise retransmit it to the make cluster on bases of EM algorithm and select cluster head process.

IV. RESULTS

Fig. 1 Drop packet vs. round number

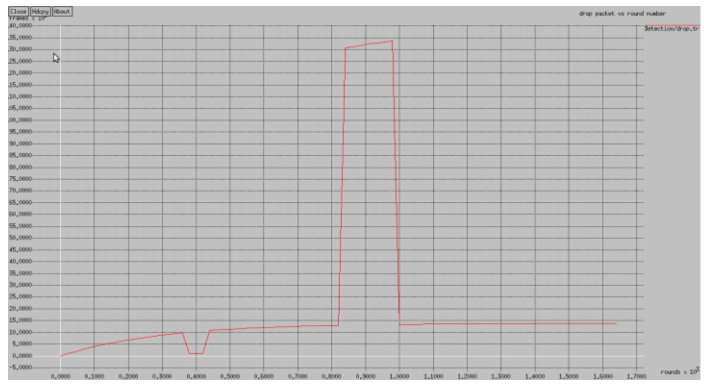


Fig.2. Dropationvs time

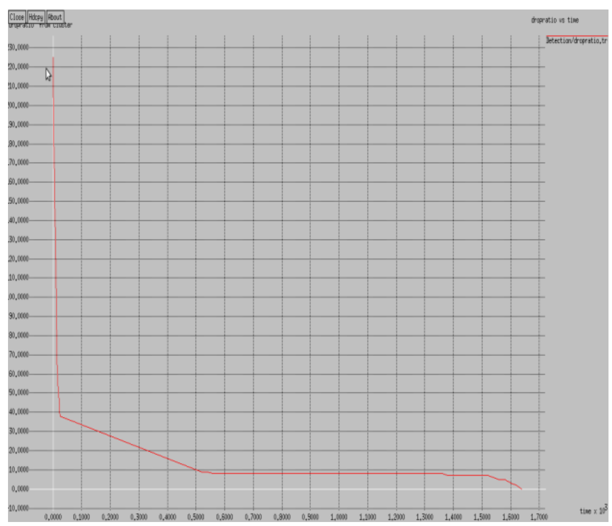


Fig. 3.end to end delay vs time

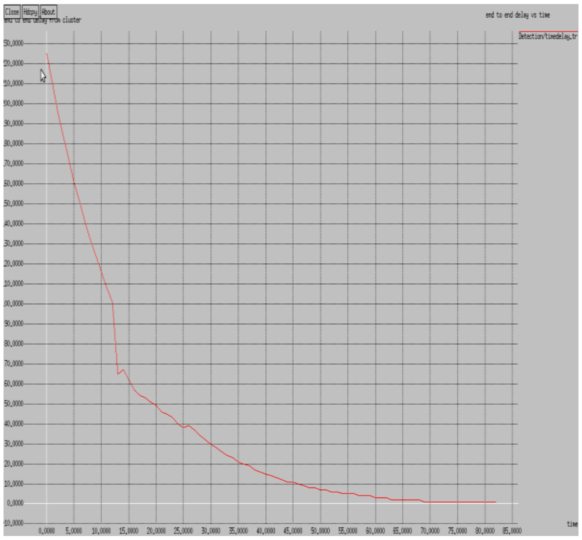
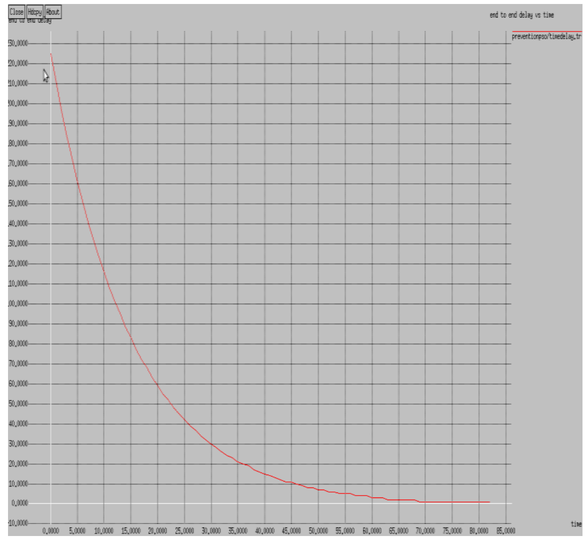


Fig.4. Throughput vs time



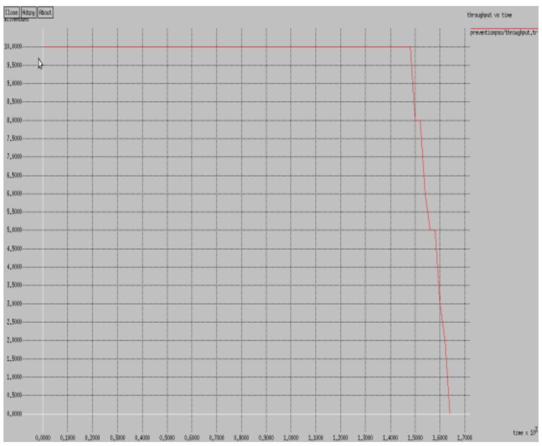


Fig. 5 drop packet vs round number

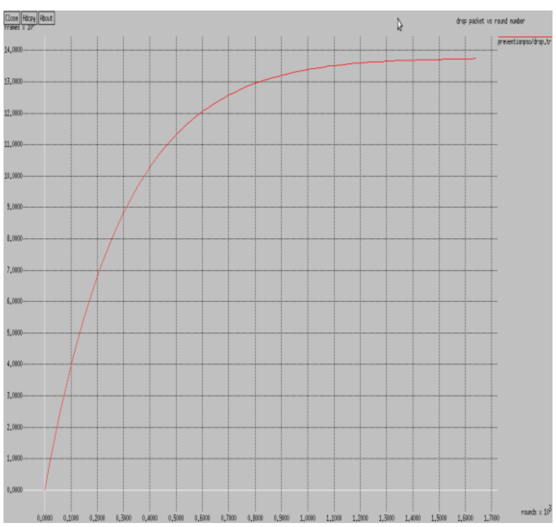


Fig. 6. End to end delay

V. REFERNCES:

[1]. Z. Ozyurt, D. Aksen, and N. Aras, "Open vehicle routing problem with time deadlines: Solution methods and an application," in Operations Research Proceedings 2005, ser. Operations Research Proceedings, H.-D. Haasis, H. Kopfer, and J. Schnberger, Eds. Springer Berlin Heidelberg, 2006, vol. 2005, pp. 73–78.

[2]. M. M. Solomon, "Algorithms for the vehicle routing and scheduling problems with time window constraints," Oper. Res., vol. 35, no. 2, pp.254–265, 1987.

[3]. Yanjun Yao, Qing Cao and Athanasios V. Vasilakos, 'EDAL: an Energy-efficient, Delay-aware, and Lifetime-balancing Data Collection Protocol for Heterogeneous Wireless Sensor Networks', 2015

[4]. Sankardas Roy, Mauro Conti, SanjeevSetia, SushilJajodia,' Secure Data Aggregation in Wireless Sensor Networks: Filtering out the Attacker's Impact', VOL. 9, NO. 4, APRIL 2014.

[5]. Thorsten Nowak, Alexander Koelpin, LucilaPatino, Falko Dressler and JornThielecke, Markus Hartmann, 'Combined Localization and Data Transmission in Energy-Constrained Wireless Sensor Networks', page(s): 6345 – 6350, 2015

[6]. BhaskarKrishnamachari, Deborah Estrin, Stephen Wicker, 'Impact of Data Aggregation in Wireless Sensor Networks', 2002. Proceedings.22nd International Conference on.IEEE, 2002.

[7]. Rajagopalan, Ramesh and Varshney, Pramod K., "Data aggregation techniques in sensor networks: A survey" (2006).

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