



To Design Host-Assisted Algorithm for Mitigation of Data Leakage in WSN using Improve Fuzzy Finger Print Technique

Kirandeep* (Student), Er. Shivali Puri (Internal Guide), Dr. Naveen Dhillon (Principal)

Department of Computer Science, RIET, Phagwara,
Punjab, India

Abstract—Wireless sensor networks (WSNs) have significant potential in many application domains such as agriculture, health, environmental monitoring, battlefield surveillance, and wild fire detection. They, however cannot be used in large geographical areas due to the short communication range of sensors.. In addition, sensor networks have been the lack of available network management and control tools, such as for determining the degree of data aggregation prior to transforming it into useful information. Designing different network management tools such for routing, localization, and data aggregation are, therefore, required in large scale WSNs. Only a few of the existing data aggregation methods have been developed for a large scale WSN. In this paper, review of some literature and techniques for data aggregation has been done in large scale WSN.

Keywords— WSN, Data Aggregation, Data Clustering and Fuzzy Fingerprint

I. INTRODUCTION

With advancement in technology, sensor networks composed of small and cost effective sensing devices equipped with wireless radio transceiver for environment monitoring have become feasible. The key advantage of using these small devices to monitor the environment is that it does not require infrastructure such as electric mains for power supply and wired lines for Internet connections to collect data, nor need human interaction while deploying. These sensor nodes can monitor the environment by collecting information from their surroundings, and work cooperatively to send the data to a base station, or sink, for analysis.

1) **Clustering in WSN:** The process of grouping the sensor nodes in a densely deployed large-scale sensor network is known as clustering.

2) **Data Aggregation:** The intelligent way to combine and compress the data belonging to a single cluster is known as data aggregation in cluster based environment.

3) **Fuzzy Fingerprint:** To achieve the privacy goal, data owner generates a special type of digests which we call fuzzy fingerprints. Intuitively, the purpose of fuzzy fingerprints is to hide the true sensitive data in a crowd. It denies the DLD (Data Leakage Detection) to access the exact value. In finger-printing, each shingle is treated as a polynomial $q(x)$. Each coefficient of $q(x)$, i.e., c_i ($0 < i < k$), is one bit in the shingle. $q(x)$ is mod by a selected irreducible polynomial $p(x)$.

II. RELATED WORK

Scalable and Unified Management And Control (SUMAC) is a large scale Wireless Sensor Network (WSN) architecture that uses a medium range mesh network as a bridge between geographically dispersed sensors clusters and Internet and provides users full data ownership and transmission of data within their own network. In SUMAC a high level setting (by users) that triggers a background process to set a default data aggregation level and also to determine rules and conditions to modify the default aggregation level is proposed. For instance, default aggregation level can be set to average most sensors data in a single cluster. When an event of interest occurs sensors can automatically slide down the aggregation level to enable node to send raw sensor data. SUMAC aggregation contains energy cost for path generation. Nodes share their energy consumptions, delay and buffer size with immediate neighbors. This method is coupled with a visual interface to control the aggregation (static or dynamic) based on the assigned aggregation rules such as hop count, resolution, sensors value, sensor tags, and/or GPS coordinate. In static aggregation, user selects nodes based on GPS coordinate, label or ID and sets the selected node to aggregate or to forward their data. Server statically sets the node aggregation level. In dynamic aggregation, aggregation level changes based on the events of interest occur. For example, if the current aggregation level is 2 and any node at level 3 detects an event of interest the aggregation level slides down to 4 so that data at level 3 are not aggregated. Users set the maximum/minimum threshold for sensed values such as temperature or light. If the temperature exceeds a defined threshold, for example, the sensor nodes automatically stop aggregating the packets and also instruct neighbors to do so. In Directed Diffusion (DD), interest messages flow from the sink to the source using expensive flooding (Interest propagation), then data messages flow from the source to the sink initially along multiple paths towards the sink (data propagation). As time progresses, the sink reinforces only a on data quality) and number of paths (depending hence the total number of nodes in transmission is reduced. To alleviate expensive flooding for interest propagation in DD, clustering approaches are used, where interest messages are only sent to cluster head (CH) and gateways. Chatterjea S. and Havinga P. propose Clustered Diffusion with Dynamic Data

Aggregation (CLUDDA) that improves energy and network efficiency by integrating clustering into Directed Diffusion (DD) and allowing nodes to collect and aggregate data by including entire query definition with interest message. The format of interest packet is significant in interest transformation, dynamic aggregation, and point formation. It also allows nodes to deal with unfamiliar queries. Interest packets contain not only the query but also the entire definition of the query. They allow nodes to break down a query into its fundamental components and gather data for these individual components and process them using the query definitions, which in turn results in data reduction. Ying Liang and Hongwei Gao propose an Optimal Clustering Algorithm Based on Target Recognition (OCABTR) that collects data periodically and hence, reduces transmission overload and energy consumptions of sensor nodes. When clusters are formed, sensors reside in different clusters might represent the same geographical area in terms of events to sense that increases data redundancy. It is also difficult to aggregate similar data in different clusters. Hence, OCABTR uses genetic algorithm to partition nearby/adjacent nodes (to form cluster) that sense similar events into a cluster that improves the rate of data aggregation. Data aggregation based on dynamic routing (DABDR) is another cluster based aggregation routing protocol. DABDR creates tree structure where parents wait a certain time for child data. Data packets have a depth field that ensures the direction of data flowing from a sampling node to sink and a queue length field that makes data packets flow to nodes with a long data aggregation queue so that data packets are concentrated more to make the aggregation more energy efficient. Tiny Aggregation Approach (TAG) is also a dynamic data aggregation method where, each epoch or time duration is divided into timeslots. Different levels of tree are associated with different timeslots and nodes of each level can only send data in their specific timeslots. Hence, synchronization is achieved for sending and receiving data that reduces energy consumptions. However, in this approach if a node does not receive the data for a child at its specified timeslot the unused information of the whole sub-tree rooted at that child will result an ultimate inaccurate data. Bidirectional Data Aggregation (BDA) adds a label to each query in addition to the basic working principle of TAG.

III. DATA AGGREGATION APPROCHES IN WSN

Data aggregation process is performed by specific routing protocol. Our aim is aggregating data to minimize the energy consumption. So sensor nodes should route packets based on the data packet content and choose the next hop in order to promote in network aggregation.

Basically routing protocol is divided by the network structure, that's why routing protocols is based on the considered approaches.

- 1) Tree-Based Approach
- 2) Cluster-Based Approach
- 3) Multi-Path Approach
- 4) Hybrid Approach

1) **Tree-Based Approach:** The tree based approach is defining aggregation from constructing an aggregation tree. The form of tree is minimum spanning tree, sink node consider as a root and source node consider as a leaves. Information flowing of data start from leaves node up to root means sink(base station). As we know like wireless sensor network are not free from failure .in case of data packet loss at any level of tree, the data will be lost not only for single level but for whole related sub tree as well. This approach is suitable for designing optimal aggregation techniques'.

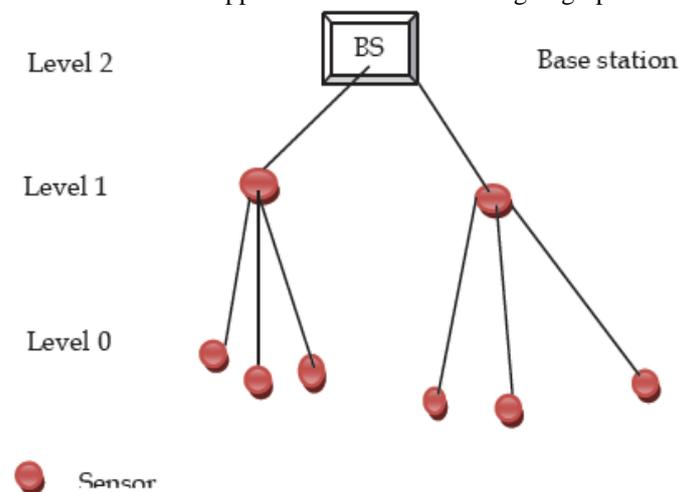


Fig 1: Tree based data aggregation

2) **Cluster-Based Approach:** In energy-constrained sensor networks of large size, it is inefficient for sensors to transmit the data directly to the sink. In such scenarios, Cluster based approach is hierarchical 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. Cluster-heads do the role of aggregator which aggregate data received from cluster members locally and then transmit the result to base station (sink). Recently, several cluster-based network organization and data-aggregation protocols have been proposed for the wireless sensor network. Figure shows a cluster-based sensor network organization.

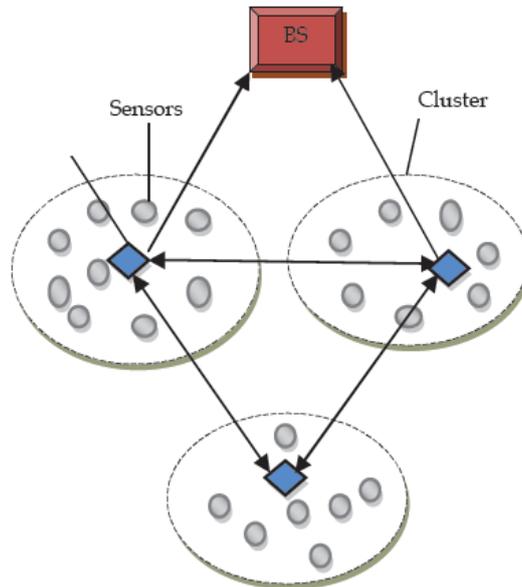


Fig 2: Cluster based data aggregation

3) **Multi-path Approach:** The drawback of tree based approach is the limited robustness of the system. To overcome this drawback, a new approach was proposed by many researchers .in which sending partially aggregated data to single parent node in aggregation tree, a node could send data over multiple paths. In which each and every node can send data packets to its possibly multiple neighbors. Hence data packet flow from source node to the sink node along multiple path, lot of intermediate node between source node to sink node so aggregation done in every intermediate node. Using this approach we will make the system robust but some extra overhead.

4) **Hybrid Approach:** Hybrid approach followed between tree, cluster based and multipath scheme. In which the data aggregation structure can adjusted according to specific network situation and to some performance statistics.

IV. HOST ASSISTED ALGORITHM

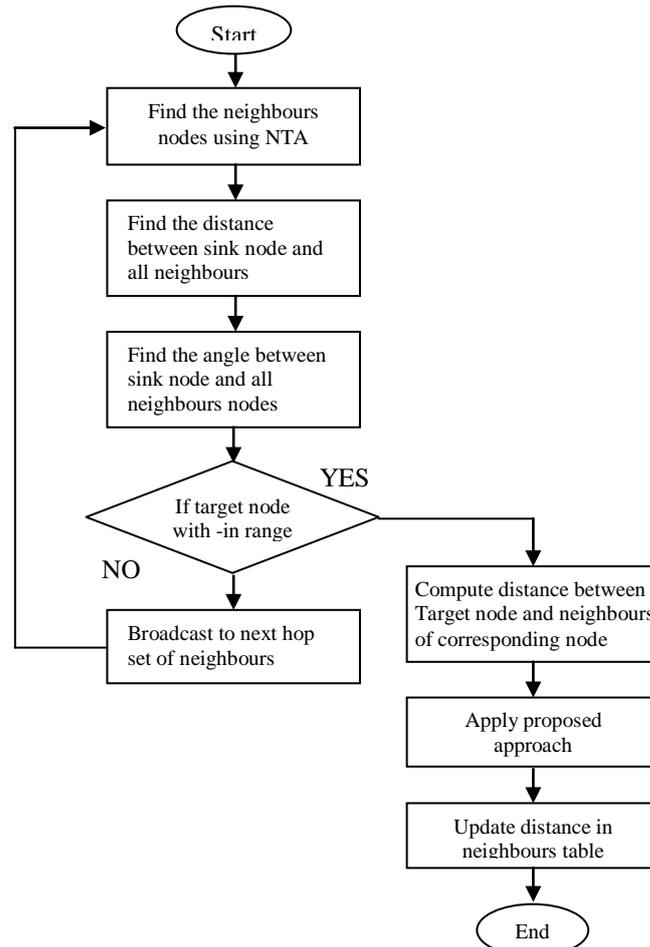


Fig 3: Flow Chart

Input: Wireless Sensor Network Nodes

Step 1: Nodes deployment over an area of 400*400

Step 2: Sink node will broadcast message in network.

Step 3: Nodes in network acknowledge sink node about the message using RSS.

Step 4: Calculate Cartesian distance for distance measurement.

Step 5: Find angle between sink node and all neighbor node.

Step 6: If target node is within NTA (Node Transmission Area)

Step 7: Source send data packet to neighbor node and update hop count that will be one greater than that of neighbor node.

Step 8: Data packet will be send to sink in same fashion.

Step 9: Calculate result against parameters like, data send, overheads, power consumption, throughput etc.

Step 9: Generate and validate results.

V. RESULTS AND DISCUSSION

In this scenario a comparison is made between hybrid routing schemes by taking 100 subscriber stations which is shown below:

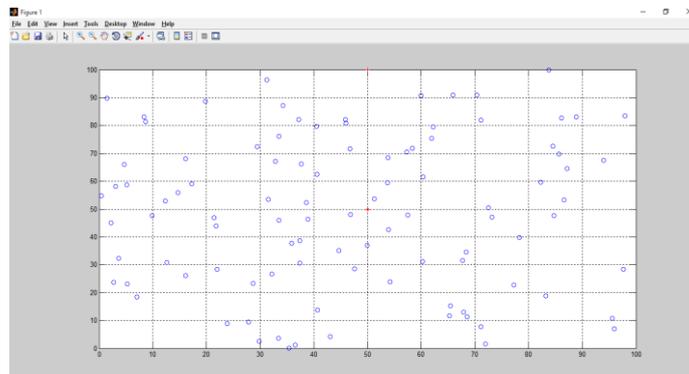


Fig 4: Node Deployment

Fig 4 is the presentation of number of nodes deployed in an area of 100*100m². In this figure there is a sink location that is specified at a location of x = 50 and y = 100 on which the packets are to be received.

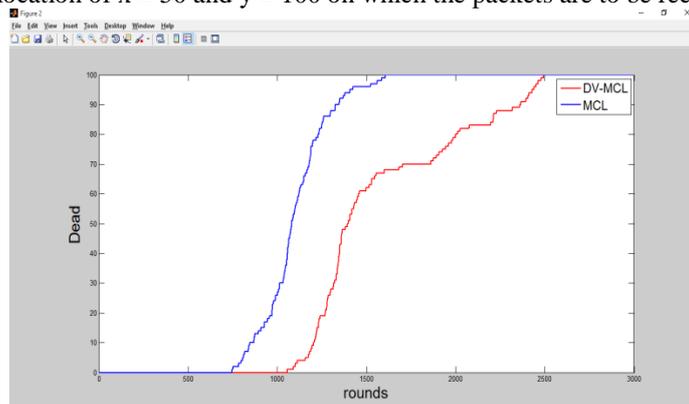


Fig 5: Dead Nodes v/s Rounds

Fig 5 is the presentation of number of dead nodes against the number of rounds. This figure is comparative study of proposed technique that is fuzzy approximation (DV-MCL) based MCL and existing technique that is MCL. From the figure it is clear that the dead nodes in proposed technique is less than that of existing technique.

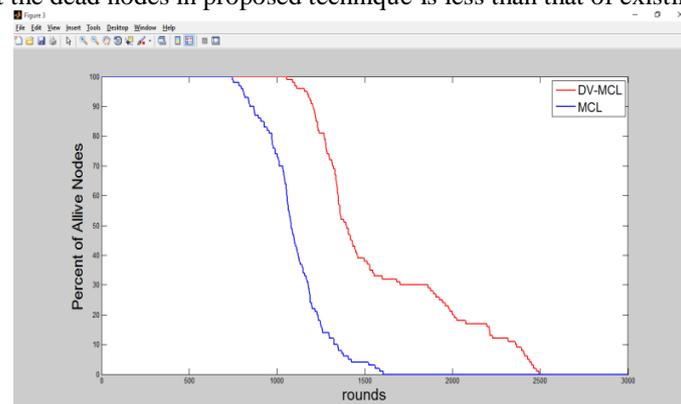


Fig 6: Percentage of Alive Nodes

Fig 6 is the presentation of percentage of alive nodes against the number of rounds. This figure is comparative study of proposed technique that is Fuzzy Approximation based MCL and existing technique that is MCL. From the figure it is clear that the alive nodes in proposed technique are less than that of existing technique.

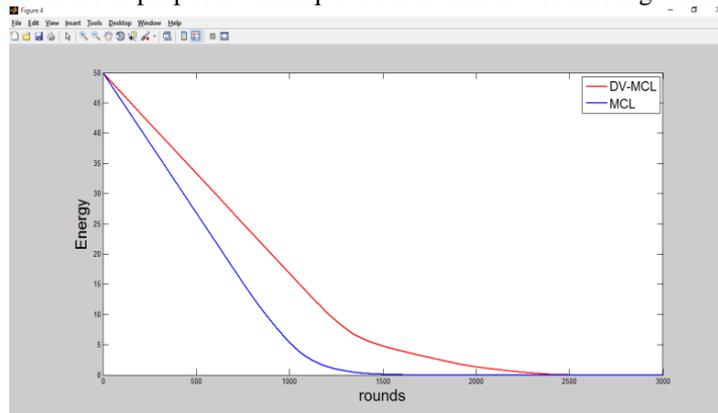


Fig 7: Energy Residual

Fig 7 is the presentation of energy residual against the number of rounds. This figure is comparative study of proposed technique that is Fuzzy Approximation based MCL and existing technique that is MCL. From the figure it is clear that the energy residual in proposed technique is more than that of existing technique.

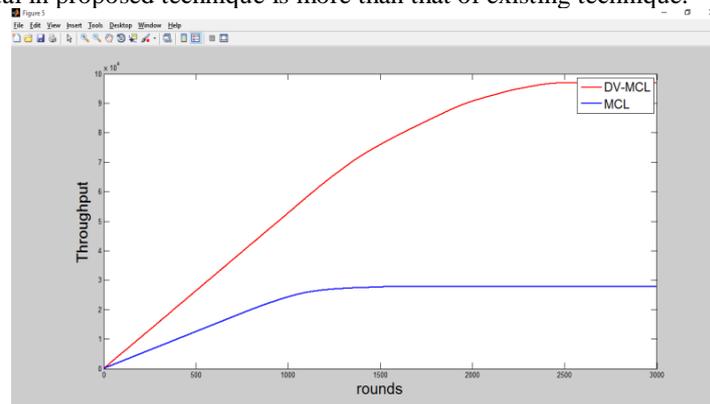


Fig 8: Throughput

Fig 8 is the presentation of throughput against the number of rounds. This figure is comparative study of proposed technique that is Fuzzy Approximation based MCL and existing technique that is MCL. From the figure it is clear that the throughput in proposed technique is more than that of existing technique.

Table 1: Comparative study of various parameters for both algorithm.

Technique Parameters	Existing	Proposed
Dead Nodes	1500 rounds	2500 rounds
Alive Nodes	1500 rounds	2500 rounds
Energy Consumption(j)	20 rounds	16 rounds
Throughput(Kb/s)	18	21

VI. CONCLUSION

This paper presents wireless sensor network is consist a large number of sensor node. And these nodes are resource constraint. That's why lifetime of the network is limited so the various approaches or protocol has been proposed for increasing the lifetime of the wireless sensor network. The paper discusses, the data aggregation is one of the important technique for enhancing the life time of the network and security issues like data integrity. With the help of integrity we reduce the compromised sensor source nodes or aggregator nodes from significantly altering the final aggregation value.

REFERENCES

- [1] Kiran Maraiya, Kamal Kant, Nitin Gupta, "Wireless Sensor Network: A Review on Data Aggregation", International Journal of Scientific & Engineering Research Volume 2, Issue 4, April -2011
- [2] Lutful Karim, Nidal Nasser, Hanady Abdulsalam, Imad Moukadem, "An Efficient Data Aggregation Approach for Large Scale Wireless Sensor Networks."

- [3] Sanjay Eknath Gawali, Prof. D.S.Mantri, “*Lifetime Energy Efficient Optimization for WSN*”
- [4] Prakashgoud Patil, Umakant P Kulkarni, “*Energy Efficient Aggregation with Divergent Sink Placement for Wireless Sensor Networks*”, International Journal of Ad hoc, Sensor & Ubiquitous Computing (IJASUC) Vol.4, No.2, April 2013
- [5] Neeraj Kumar Mishra, Vikram Jain, Sandeep Sahu , “ *Survey on Recent Clustering Algorithms in Wireless Sensor Networks* ”, International Journal of Scientific and Research Publications, Volume 3, Issue 4, April 2013
- [6] Karim Seada, Marco Zuniga, Ahmed Helmy, Bhaskar Krishnamachari, “*Energy Efficient Forwarding Strategies for Geographic Routing in Lossy Wireless Sensor Networks*”
- [7] Xun Li, Geoff V Merrett, Neil M White, “*Energy-efficient data acquisition for accurate signal estimation in wireless sensor networks*”, Journal on Wireless Communications and Networking 2013
- [8] Gyanendra Prasad Joshi, Seung Yeob Nam and Sung Won Kim, “*Cognitive Radio Wireless Sensor Networks: Applications, Challenges and Research Trends.*”
- [9] Stefanos A. Nikolidakis, Dionisis Kandris, Dimitrios D. Vergados, Christos Douligeris, “*Energy Efficient Routing in Wireless Sensor Networks Through Balanced Clustering*”
- [10] Wei Ye, John Heidemann, Deborah Estrin, “*An Energy-Efficient MAC Protocol for Wireless Sensor Networks*”
- [11] Mohit Saini, Rakesh Kumar Saini, “*Solution of Energy-Efficiency of sensor nodes in Wireless sensor Networks*”, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 5, May 2013
- [12] Mark Adam Perillo, “*Role Assignment in Wireless Sensor Networks: Energy-Efficient Strategies and Algorithms*”
- [13] Shashidhar Rao Gandham, Milind Dawande, Ravi Prakash, S. Venkatesan, “*Energy Efficient Schemes for Wireless Sensor Networks with Multiple Mobile Base Stations.*”
- [14] Eugene Shih, Seong Hwan Cho, Nathan Ickes, Rex Min, Amit Sinha, Alice Wang, Anantha Chandrakasan, “*Physical Layer Driven Protocol and Algorithm Design for Energy Efficient Wireless Sensor Networks*”
- [15] N. Akilandeswari, B. Santhi and B. Baranidharan, “*A Survey on Energy Conservation Techniques in Wireless Sensor Networks*”, ARPN Journal of Engineering and Applied Sciences, VOL. 8, NO. 4, APRIL 2013
- [16] Neeraj Kumar, Manoj Kumar, and R. B. Patel, “*A Secure and Energy Efficient Data Dissemination Protocol for Wireless Sensor Networks*”, International Journal of Network Security, Vol.15, No.6, PP.490-500, Nov. 2012