
Sonal Sharma  
Department of 
Computer Science and Engineering  
Institute of Technology & Management  
Gwalior - INDIA

Rajendra Singh Kushwah  
Department of 
Computer Science and Engineering  
Institute of Technology & Management  
Gwalior - INDIA

ABSTRACT

Wireless sensor networks (WSNs) incorporate sensor nodes. These networks have gigantic application in habitat monitoring, disaster management, protection and military, etc. Wireless sensor nodes are very small in measurement and have limited processing ability very low battery power. This restriction of low battery power makes the sensor network susceptible to failure. Data aggregation could be very imperative process in WSNs. Data aggregation is new method for energy consumption in WSN. This paper is expending greedy approach and ant colony optimization technique in Data aggregation.

Keywords

WSNs, Data Aggregation, Greedy Algorithm, Ant Colony Optimization (ACO)

1. INTRODUCTION

WSNs are for the most part perceived as another innovation which consolidates countless wireless sensor nodes with spatial distribution to bolster a wide variety of uses, including indigenous environment monitoring, medical administrations, observation and ocean pollution detection. In an extensive scale proactive WSN, each sensor node performs occasionally a few operations proportionate to figuring, detecting and self-arranging to transmit precise data to the sink node through different ways. In this kind of design, sensors are utilized and substantial fueled by method for limited lifetime batteries, which can be hard to be replaced or revived. Distinctive resource restrictions in WSNs are short correspondence range, low transfer speed, restrained processing / stockpiling and itemized the energy usage. Energy usage is generally tended to in the accompanying three phases: detecting, data get ready, and data transmission. For the most part, detecting and data handling have less energy utilizations than that of data transmission. For sure, any decrease in the transmission cost can drag out the WSN's lifetime. As an outcome, minimizing the complete energy utilization is of the top worth in outlining WSNs. Various exploration works have tended to the energy productivity wander in WSNs from stand-out perspectives, including energy scam serving rest planning, topology oversee , mobile data powers , and data aggregation. unsafe to this learn is to set up legitimate data aggregation and routing strategies in a WSN to construct every energy use and the network’s lifetime with taking the consequence of load balancing under thought.

Data aggregation has been advanced as a crucial worldview for wireless routing in sensor networks. The contemplation is to combine the data beginning from unprecedented sources enrooted – disposing of redundancy, minimizing the amount of transmissions and consequently sparing energy. This viewpoint changes the motivation behind eagerness from the normal handle-driven procedures for systems networking (finding short route between sets of addressable end-nodes) to an extra additional data-method (finding routes from different sources to a single destination that grants in-network blend of repetitive data).

2. LITERATURE REVIEW

Data aggregation algorithm means to collect information or gathered information in an very efficient manner, so that network access availability will become large. WSN offers an increasingly nodes that needs the arability in less amount for processing or for transmitting data from source to destination place.

Different aggregation features would be max, min, or another function with a couple of inputs. For our modeling capacities we make an improving assumption - the aggregation capacity is handled in a manner that each intermediate node inside the routing transmits a solitary aggregate packet regardless of the possibility that it gets different input packets. We will imply the data got by the sink when it has gotten the messages transmitted by all sources in a given stream (regardless of whether these messages are aggregated) as a “datum”.

2.1 Optimal Aggregation

Say there are k sources, named S1 through Sk, and a sink, named D. Give the network graph G = (V,E) incorporate the greater part of the nodes V , with E which incorporate edges between nodes that may keep up a correspondence with each distinctive straight. With the suspicion that the amount of transmissions from any node inside the data aggregation tree is accurately one, the data aggregation tree can likewise be considered as the inverse of a multicast tree: as opposed to a single supply sending a packet to all beneficiaries, every one of the sources are sending a singular packet to the same collector. It is comprehended that the multicast tree with a base number of edges is a base Steiner tree on the network graph. The accompanying can thusly be effectively gotten:

Result 1: The ideal number of transmissions required per datum for the DC protocol is the same as the amount of edges inside the base Steiner tree in the network which includes the node set (S1, ... Sk, D).

Corollary: Assuming a self-assertive arrangement of sources, and a typical network graph G, the task of doing DC protocol with most helpful data aggregation is NP-hard. The last takes after from the NP-satisfaction of the base Steiner inconvenience on Graphs.

2.2 Suboptimal Aggregation

The following are three regularly imperfect plans for producing data aggregation trees that we assess on this paper.
2.2.1 Center at Nearest Source (CNS)
On this data aggregation plot, the source which is nearest the sink exhibits in light of the way that the aggregation point. Every single assorted source sends their data instantly to this source which then sends the aggregated data on to the sink.

2.2.2 Shortest Path Tree (SPT)
On this data aggregation conspire, every last source sends its data to the sink close by the briefest route between two. This way every node cover for extraordinary sources, they are blended to sort the aggregation tree.

2.2.3 Greedy Incremental Tree (GIT)
In this arrangement the aggregation tree is produced progressively. At the underlying stride the tree consolidates only the most limited way between the sink and the nearest source. At every last stride after that the following supply closest to the current tree is associated with the tree. This is by no means, an extensive summary, yet is illustrative of a portion of the data aggregation tree heuristics that can be executed.

2.3 Performance Measure
In investigating the increases and tradeoffs required in data centric protocols, we have to indicate execution measures of premium. Two are inspected in some point of interest in this paper:

2.3.1 Energy Savings
By aggregating the data originating from the sources, the quantity of transmissions is reduced, which means a store in energy [3].

2.3.2 Delay
There is inertia associated with aggregation. Data from nearer sources may kept down at moderate nodes to join them with data from sources those are more distant away.

2.4 Source Placement Model
The official reasons that may influence the efficiency of data aggregation methodologies are the positions of the sources in the network, the amount of sources and correspondence network topology. With a particular finished objective to investigate these segments, this paper focus on two models of source course of action, the event-radius (ER) model, and the random sources (RS) model.

In each models, we make a sensor network through dissipating n sensor nodes randomly in a unit square. All nodes are thought to be prepared to keep up a correspondence with other nodes which can be inside far R (the communique radius). The spot of the data sources is controlled by the things as takes after.

2.4.1 Event Radius Model
In this model, a solitary point in the unit rectangular is characterized in light of the fact that the region of an “event” this may relate to an auto or some other wonder being followed with the guide of the sensor nodes. All nodes inside a separation S (known as the detecting extent) of this occasion are not sinks are thought to be data sources.

2.4.2 Random Source Model
In this model, nodes that are not sink are randomly been source. Not in any manner like in the occasion range demonstrate, the sources are not as a matter of route clustered near to each other.

3. DATA AGGREGATION IN WSN
Data Aggregation is an arrangement of joining and outlining the data from sensor nodes in WSNs through making utilization of aggregation capacity relating to MAX, MIN, AVG, depend, SUM as in etcetera on aggregator nodes. Data Aggregation is an arrangement of expelling repetitive information from different sensor nodes. Data aggregation frameworks as in characterized approach define how the Data is to be directed on the network and handling technique that are used on the Data packets.

3.1 Data Aggregation Approaches
There are more than a couple of techniques of data aggregation some are as per the following:

3.1.1 Centralized Approach
In this methodology one and only sensor node have impact of aggregator node and all other sensor nodes are associated with that aggregator node. All distinctive sensor nodes experience the data and transmit to the aggregator node which is called brought concentrated node. There is such a broad measure of weights on that aggregator node, so there might be need of more vitality and security on that aggregator node in light of the way that all learning is on the unified aggregator node.

3.1.2 Decentralized Approach
In this approach all sensor nodes performs aggregator work to the sensed data. There is no centralized aggregator node yet all nodes have same requirement to aggregate the sensed data. In this approach all sensor nodes are identified with their neighbor node. This approach has benefit of additional scalability, dynamic changes node failure in the WSN.

3.1.3 In Network Aggregation Approach
In this methodology different node will in like manner be aggregator node procedure sub-aggregator node. This methodology aggregates numerous data into single data. It is essential for enhancing the network lifetime and diminishes the span of transmitted data on the network

4. TABLE 1: COMPARATIVE STUDY OF DATA AGGREGATION ALGORITHM

<table>
<thead>
<tr>
<th>Topic</th>
<th>Technique</th>
<th>Latency</th>
<th>Energy</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatio-temporal correlation</td>
<td>Cluster based</td>
<td>Yes</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Scale Free Aggregation</td>
<td>Cluster based</td>
<td>No</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Impact of Network Density on Data Aggregation</td>
<td>Tree based</td>
<td>No</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>An Ultra-Low power And Distributed Access Protocol</td>
<td>Cluster based</td>
<td>Yes</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Interdependence of Routing</td>
<td>Multihop based</td>
<td>No</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>and Data Compression</td>
<td>Cluster based</td>
<td>Yes</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------</td>
<td>-----</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Spatio-Temporal Sampling Rates</td>
<td>Yes</td>
<td>No</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Spatial and Temporal Multi-Aggregation</td>
<td>No</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Hybrid, Energy-Efficient, Distributed Clustering</td>
<td>Cluster based</td>
<td>No</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Polynomial regression for data representation</td>
<td>Tree based</td>
<td>Yes</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td>Redundancy Elimination for Accurate Data Aggregation</td>
<td>No</td>
<td>Low</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>An Energy Efficient Clustering Scheme for Data Aggregation</td>
<td>No</td>
<td>High</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td>An Energy Efficient Spatial Correlation Based Data Gathering</td>
<td>Clustering based</td>
<td>No</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Energy Efficient and Balancing</td>
<td>Cluster based</td>
<td>No</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

Data aggregation structure alongside the algorithm on WSNs is exhibited. The purpose of the proposed work includes a nitty-gritty study with respect to the mistakes which can happen by the data aggregation, furthermore the circuit multifaceted nature which is generated for reduction of these errors [6]. Evaluation results verify the effectively of our scheme in terms of the performed end-to-end transmission delay for periodic packets and the diminished overhead. These result uncover furthermore that our plan beats diverse three literature schemes (no aggregation, randomized prepared and full aggregation) since it ensures the best exchange off for aggregation sparing and delay reduction [7]. A few security issues which incorporate data integrity, data classification, availability, and freshness in data aggregation that get to be basic when WSN is sent in an hostile environment the spot sensors fall slanted to node screw ups and traded off through foe. As an outcome more than a couple assurance issues in aggregation will probably be examined [8]. Distinguish the adjoining join planning issue in WSNs, where every node is relegated back to back time spaces so that the node can wake up only once in a booking period to satisfy its data assortment task. The goal of the impediment is to discover an obstruction free hyperlink scheduling with the base number of time spaces utilized [9]. Handle the data aggregation and assurance issues together. In our methodology, we modify our Energy efficient Cluster Based data Aggregation (ECBDA) plan to outfit secure data transmission. When you consider that, sensors nodes are low fueled in nature, it’s not possible to utilize standard cryptography approaches. Cluster head performs data aggregation and Bayesian fusion algorithm to allow insurance [10]. Differential data transfer accomplishes more energy and bandwidth affectivity than the current recoverable concealed data aggregation plan. The proposed differential data set up recoverable data aggregation plan increases the network lifetime by means of keeping off the redundant data transfer from each and every sensor node [11]. Focus on the traded off node and denial-of-service is the two key attacks in WSNs. We differ that multipath routing systems are absolutely powerless to such attacks. In this way, for these attacks we improve the components that produce randomized multipath routes. In this planning, the routes are taken by method for the shares of different packets transform after some time [12] they reward an energy effective, privacy safeguarding data aggregation algorithm which also preserves data integrity in WSNs. We break down the insurance of the algorithm and outfit proofs for classification and integrity [13]. Orchestrated on iritation based compelling protection keeping arrangement and Energy-Efficient Secure Pattern Based Data Aggregation for WSNs arrangement, we obtain greater power-efficiency with the aid of utilizing pattern-codes procedure to restrict energy wastage because of sending redundant data to cluster-heads [14]. Data aggregation algorithm on Mica2 and TelosB sensor network structure measures the execution time and energy usage of more than two or three cryptographic capacities. We now have also reproduced our algorithms to examine how a completion to end plan expands the network life time. We tentatively analyze our algorithms established on parameters like throughput, end to finish length and resilience to node failures [15] energy efficient algorithm for gathering data from all sensor nodes with slightest idleness insinuated as deferral minimized Energy-effective data Aggregation algorithm (DEDA). The DEDA algorithm minimizes data aggregation torpidity by strategy for building up a deferral successful network structure [16] an overview on data aggregation framework in WSN. Here, we find particular types of data aggregation algorithms arranged on network structure, frameworks and execution attributes. In light of the study, the future examination territories and issues to be engaged are Security, Energy protection, question preparing, correspondence and instability in sensor reading [17]. Secure data aggregation protocol for WSNs is staggering to beguiling nodes. The purpose of this protocol is to guarantee the foremost protection desires (like source authentication, data confidentiality & data integrity) as good as to obtain low communication overhead and be equipped with more than a few aggregation capabilities (like sum, natural, max, min etc.)[18]. To gain these safety needs, it uses symmetric encryption and Message Authentication Code (MAC). Encryption ensures information protection while message verification code guarantees confirmation and data integrity. A peculiarity detection algorithm is utilized to watch the oddity or exceptions and thus turn away the tricky adulterated information from being added to the last aggregated outcome. Simulation results exhibit that our protocol enhances the safety of the aggregated data considerably in WSNs [18].
5. GREEDY APPROACH
An greedy algorithm is a algorithm that used to discover the ideal arrangement at every phase which takes after the critical thinking heuristics. It is also an numerical procedure that searches for simple, easy to execute answer for complex, multistep issues by choosing which next stride will give the most obvious benefits. General guideline of greedy algorithm movement choice issue this issue likewise shows the greedy property.

If we apply greedy approach in WSN than there are two most important issues of sensor network one is energy of nodes and second is security we concern about energy of nodes, so that for cluster head selection we choose greedy approach.

If ((energy is high && (density > threshold) && (movability > slow)) { 
    Become cluster head. }

6. ANT COLONY OPTIMIZATION
(ACO)
The ant colony algorithm is a algorithm for finding ideal ways that is set up on the behavior of ants searching for food. In any case, the ants wind self-assertively. Right when an insect finds a source of food, it walks around to the state leaving “markers” (pheromones) that exhibit the way has food. Exactly when distinctive ants come all through the markers, they’re inclined to bring after the path with a nitty itemized likelihood. In the event that they do, they then populate the way with their own markers as they bring the food back. As more ants discover the way, it will get predominant with the exception of there are a few floods of ants setting out to different food sources close to the colony. Considering the way that the ants drop pheromones whenever when they convey nourishment, shorter ways will probably be more reasonable, in this manner optimizing the “answer.” In the interim, a few ants are still haphazardly scouting for nearer food sources. An indistinguishable system can be used to find close optimal answer for the voyaging sales representative impediment. Once the food supply is depleted, the route is not any more populated with pheromones and gradually rots. For the reason that the subterranean ant colony works takes a shot at an astounding dynamic philosophy, the ant colony algorithm works amazingly well in charts with evolving topologies. Instance of such systems include computer networks and counterfeit consciousness reenactments of delegates [20].

ACO:
procedure
ACOMetaheuristic
ScheduleActivities
ConstructAntsSolutions
UpdatePheromones
DaemonActionS
% optional end-Schedule Activities
end-procedure

7. CONCLUSION
WSNs consist of sensor nodes. These networks have huge utility in habitat monitoring, disaster management, protection and military, and so forth. Wireless sensor nodes are very small in size and have constrained processing potential and really low battery vigor. Data aggregation helps in diminishing the vitality utilization by eliminating redundancy. This paper includes greedy approach and ant colony optimization technique in data aggregation. For cluster head selection greedy approach i.e. cost effective and time saving approach is used, so that proposed approach is more effective, optimization technique is used for a routing i.e. gives better effective routing in WSN. This work gives better result in form of network lifetime and energy of nodes decrease in constant manor, proposed approach is also effective in form of data gathering.

8. REFERENCES
[6] Imane Horiya Brahmi, SoufieneDjahel, Damien Magoni z and John Murphy “A Spatial Correlation Aware Scheme for Efficient Data Aggregation in Wireless Sensor Networks” LCN 2015, Clearwater Beach, Florida, USA.


