ISSN: 2277-9655
Impact Factor: 1.852

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Data Aggregation Using Genetic Algorithm in Wireless Sensor Network

Priti Saluja[*1], Sandeep Dalal[2]

[*1] M.Tech (CS), Department of Computer Science and Applications, MDU, Rohtak (Haryana-124001), India

[2] Assistant Prof., Department of Computer Science and Applications, MDU, Rohtak (Haryana-124001), India

pritisaluja.85@gmail.com

Abstract

A sensor network consists of one or more “sinks”. The sensors in the network act as “sources” which detect environmental events and push relevant data to the appropriate sinks. Sensors transmit information towards the sink if and when they detect the indicated phenomenon. Due to the redundancy present in the sensors’ readings, it is expected that communication approaches that take into account this redundancy, e.g., data aggregation and in-network processing. The main idea of the data aggregation and in-network processing approaches is to combine the data arriving from different sources (sensor nodes) at certain aggregation points (or simply aggregators) eliminate redundancies by performing simple processing at the aggregation points, and minimize the total amount of data transmission before forwarding data to the external BS (base station or sink).Genetic algorithm (GA) is used to create energy efficient data aggregation trees. During a chromosome, the gene index determines the node and the gene’s value identifies the parent node. Single-point crossover and mutation operators are used to create future generations. In this research work, by using various mutation of genetic algorithm a new algorithm is proposed for data aggregation in WSN.

Keywords: WSN (Wireless Sensor Network), Data Aggregation, Genetic Algorithm, Mutation.

Introduction

Wireless sensor networks consist of sensor nodes with sensing and communication capabilities.WSN are highly distributed networks of autonomous small, lightweight sensors (nodes) in large numbers to monitor physical or environmental conditions by the measurement ofsound, pressure, temperature, vibration, motion or pollutants and to cooperatively pass their data through the network to a mainlocation (often called a sink) [1]. A Wireless Sensor Network (WSN) consists of a large set of sensor nodes that cooperate to monitor environmental conditions (e.g., temperature, precipitation, and radio-activity) in a given geographic area. WSNs are often designed for long term operation in remote unattended environments, despite the limited battery capacity of the wireless sensor nodes. Since data transmission is an energy-intensive task, energy aware data gathering techniques are used to extend the lifetime of the WSN. An effective way to conserve energy is to avoid reporting redundant data that occurs due to the spatial correlation between nearby readings. However, minimum cost network correlated data gathering is NP-complete [10].

Data Aggregation

Sensors could be scattered randomly in harsh environments like as a battlefield or deterministically placed at specified locations. Sensors coordinate among themselves to form a communication network such as a single multi-hop network or a hierarchical organization with several clusters and cluster heads. Sensors periodically sense the data, then process it and transmit it to the base station. The frequency of data reporting and the number of sensors which report data usually depends on the specific application [6]. Data gathering is defined as the systematic collection of sensed data from multiple sensors to be eventually transmitted to the base station for processing. Ever since, sensor nodes are energy constrainedso it is inefficient for all the sensors to transmit the data directly to the base station. The data generated from neighboring sensors is often redundant and highly correlated. More, the amount of data generated in large sensor networks is usually enormous for the base station to process. In meanwhile, we need methods for combining data into high quality information at the sensors or intermediate nodes which can reduce the number of packets transmitted to the base

ISSN: 2277-9655
Impact Factor: 1.852

station resulting in conservation of energy and bandwidth. That can be accomplished by data aggregation.

Data aggregation is defined as the process of aggregating the data from multiple sensors to eliminate redundant transmission and provide fused information to the base station. Usually data aggregation involves the fusion of data from multiple sensors at intermediate nodes and transmission of the aggregated data to the base station (sink). It attempts to collect the most critical data from the sensors and make it available to the sink in an energy efficient manner with minimum data latency.

Many types of data aggregation techniques are present some of them are listed below:

Tree-Based Approach[9]

In the tree-based approach perform aggregation by constructing an aggregation tree which could be a minimum spanning tree, source nodes and rooted at sink are considered as leaves. Every node has a parent node to forward its data. The flow of data starts from leaves nodes up to the sink and therein the aggregation done by parent nodes.

Centralized Approach

This is an address centric approach where each node sends data to a central node via the shortest possible route using a multihop wireless protocol. Sensor nodes simply send the data packets to a leader that is the powerful node. The leader aggregates the data which can be queried. Every intermediate node has to send the data packets addressed to leader from the child nodes. Thus 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.

Cluster-Based Approach[8]

In this, whole network is divided in to several clusters. Particular cluster has a cluster-head which is selected among cluster members. The cluster-heads do the role of aggregator which aggregate data received from cluster members locally and then transmit the result to sink.

Grid-Based Data Aggregation

Vaidhyanathan *et. al.* [12] have proposed two data-aggregation schemes that are based on dividing the region monitored by sensor network into several grids. These are as: Grid-based data aggregation and In-network dataaggregation. In grid-based data aggregation, it contains a set of sensors is assigned as data aggregators infixed regions of the sensor network. Sensors in a particular grid transmit the data directly to the data aggregator of that grid. Therefore, the sensors within a grid do not communicate with each other. And In-network aggregation is similar to grid-based data aggregation with two major differences, namely, each sensor within a grid communicates with its neighboring sensors. Each sensor node within a grid can assume the role of a data aggregator in terms of rounds until the last node dies.

In-Network Aggregation [4]

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 for increasing network lifetime. Mainly two approaches for in-network aggregation are: with size reduction and without size reduction. In this with size reduction refers to the process of combining, merging and 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.

Genetic Algorithm Used For Data Aggregation in WSN

Data gathering is a common but it is significant operation in many applications of WSNs, while data aggregation and hierarchical mechanism are widely used techniques. The data aggregation can eliminate data redundancy and reduce communication load. Genetic Algorithm (GA) is a technique for randomized search and optimization and has been applied to a wide range of studies [3,5]. A basic operation flow of GA includes creating initial population, evaluating fitness, selection, crossover, mutation, updating optimal chromosome, and checking termination condition. Searching for optimal clusters can be done using GA. Following, are the energy-efficient clustering technique using GA that has been developed previously. Electing a cluster-head that can minimize the maximum intra-cluster distance between itself and its cluster members by optimize energy consumption of the network [11].

Proposed Work

A sensor network is composed of a large number of sensor nodes which are densely deployed either inside the phenomenon or very close to it [7]. One of the most important features in WSNs belongs to the limited battery of sensor nodes. When battery-powered wireless sensor nodes are placed in a specific field, it is complex to replace their batteries or provide additional energy. Furthermore, if one sensor node consumes completely its energy, part of the network may disconnect [13].

Data gathering is a common but it is significant operation in many applications of WSNs, while data aggregation and hierarchical mechanism are widely used techniques. The aggregation can eliminate data redundancy and reduce communication load. Data mining technique like clustering mechanisms is an

ISSN: 2277-9655
Impact Factor: 1.852

effective means for running such high population of nodes and can help reduce the nodes' energy consumption [2].

For our proposed algorithm, firstly generate the population; evaluate the fitness for all population. By using the fitness values make the clusters of the populations. Then calculate the distance between the center node and all different nodes select the cluster head as parent node from every cluster. Then apply the crossover and mutation and calculate the fitness of population.

By using genetic algorithm we proposed our algorithm for improving the performance of data aggregation:

Step 1: Generate initial population.

Step 2: Evaluate fitness of all the individuals in the population.

Step 3: Cluster the population according to the fitness value:

i) Randomly generate k cluster centers(CH)(Here k=2).
ii) Calculate distance between the center and other nodes/chromosomes in the population.
iii) Move the CH(centroid) of cluster to perform re-clustering.

Step 4: Select the CH as the parent node from each cluster.

Step 5: Apply genetic operator that means crossover and mutation.

Step 6: Here two mutation swap mutation and insertion mutation is applied.

Step 6: Calculate fitness of all individuals in the population.

Step 7: Go to step 3 to 6 until the termination criterion is satisfied.

Step 8: Transfer sensed data from each member to CH, where
$D= (d_1,d_2,\ldots d_i)$

Step 9: Perform data aggregation in each CH, using aggregate function.

Step 10: Calculate
$avg = \frac{\sum_{i=1}^{n}(d_i)}{N}$ and max = max(d_i) *and* min=min(d_i).

Results

The proposed algorithm is simulated using the Matlab 2010. The purpose of the genetic algorithm is to minimize the fitness function. Here the fitness function is the distance. It means the proposed algorithm minimize the distance. Various screens that show the initial node position and their movement with their final position are shown. A graph between the average fitness function and the fitness function is drawn to analyze the results. The packet delivery ratio as well as the end to end delay is also analyzed.

In figure 1, the cluster formation is build using genetic algorithm.

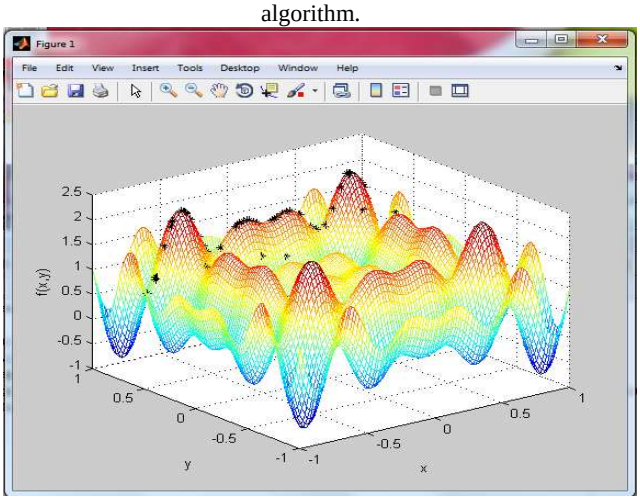

Figure: 1 Cluster formation using genetic algorithm

In figure 2, it shows the final position of the nodes.

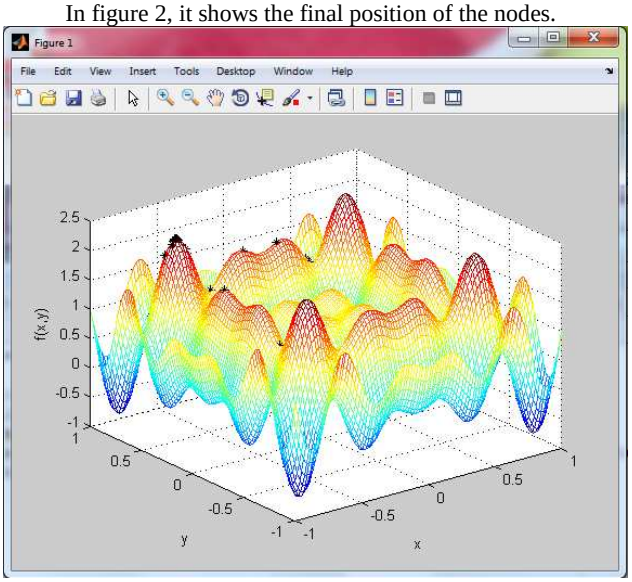

Figure2: Final position of nodes

In figure:3, the plot the maximum fitness function of the proposed algorithm is denoted by the red line and it is less than the average fitness function denoted by the blue line.

ISSN: 2277-9655
Impact Factor: 1.852

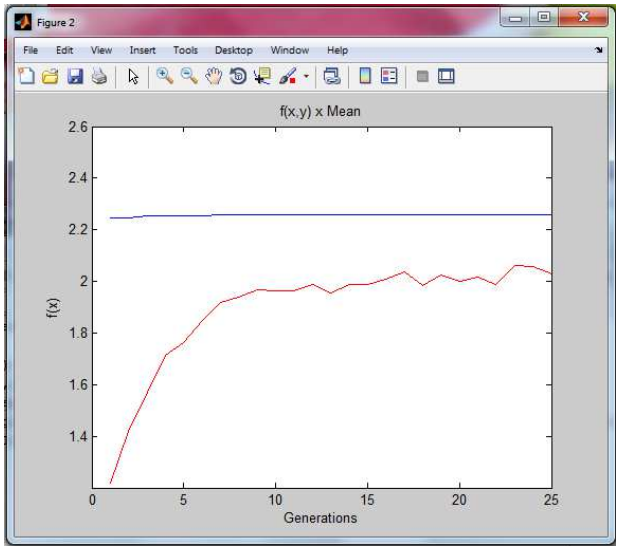

Figure3:10Max fitness function vs average fitness function

Conclusion

In WSN, the main goal of data aggregation algorithms is to gather and aggregate data in an energyefficient manner so that network lifetime is enhanced. Wireless sensor networks (WSN) offer an increasingly sensor nodes need less power for processing as compared to transmitting data. Wireless sensor networks have limited computational power and limited memory and battery power that leads to increased complexity for application developers and often results in applications that are closely coupled with network protocols**.** Given that the communication cost is several orders of magnitude higher than the computation cost so directed diffusion can achieve significant energy savings with in-network data aggregation. This benefit of data aggregation has been confirmed theoretically and experimentally. Paths from different sources to a sink form an aggregation tree rooted at the sink. Data from different sources is opportunistically aggregated. Data aggregation is preferable to do in network processing inside network and reduce packet size. By using the concept of various mutations we proposed a new algorithm for data aggregation. The simulation results in PDR(Packet delivery ration) 99% and the end to end delay is 2.03ms. Here the PDR is increase. It also reduces the end to end delay.

References

[1] Asif Khan, IsrafilTamim, Emdad Ahmed, Muhammad Abdul Awal, *"Multiple Parameter Based Clustering (MPC): Prospective Analysis for Effective Clustering in Wireless Sensor Network (WSN) Using K-Means Algorithm"*,Wireless Sensor Network, 2012, 4, 18-24 doi:10.4236/wsn.2012.41003 Published Online January 2012 (http://www.SciRP.org/journal/wsn).

[2] A. Abbasi and M. Younis, *"A Survey On Clustering Algorithms For Wireless Sensor Networks,"*Computer Communications, vol. 30, pp.2826-2841, October 2007.

[3] D. E. Goldberg, Genetic Algorithms in Search, *Optimization, and Machine Learning,Addison* Wesley, Reading, MA, 1989.

[4] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, *"In- Network Aggregation Techniques For Wireless Sensor Networks: A Survey"*,IEEE Wireless communication 2007.

[5] K. Sastry, D. Goldberg, and G. Kendall, "*Genetic Algorithms.Chapter 4 of Introductory Tutorials in Optimization and Decision Support Techniques"*, (eds. E. Burke and G. Kendall), pp. 97–125, Kluwer, 2005.

[6] I. Akyldiz, W.Su, Y. Sankarasubramanian and E. Cayirci, *"A Survey on Sensor Networks,"* IEEE Commun. Mag., vol. 40, no. 8, Aug. 2002, pp. 102-14.

[7] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "*Wireless Sensor Networks: A Survey,*" Computer Networks, vol. 38, no.4, pp. 393–422, March 2002.

[8] K. Dasgupta, K. Kalpakis, and P. Namjoshi, *"An Efficient Clustering-Based Heuristic For Data Gathering And Aggregation In Sensor Networks"*, IEEE 2003.

[9] M. Lee and V.W.S. Wong, *"An Energy-Aware Spanning Tree Algorithm For Data Aggregation In Wireless Sensor Networks,"IEEE PacRrim 2005*,

[10] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, *"On Network Correlated Data Gathering,"* in *IEEE INFOCOM*, March 2004, pp. 2571–2582.

[11] R. Khanna, H. Liu, and H. H. Chen, *"Self-Organization Of Sensor Networks Using Genetic Algorithm,"* IEEE ICC '06, pp.3377-3382,June 2006.

[12] Vaidhyanathan K. et al. *Technical Report, OSU-CISRC-11/04-TR60, Ohio State University*,2004.

[13] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan,*"Energy-Efficient Communication Protocol For Wireless Microsensor Networks,"* in Proc. of the 33rd Annual Hawaii International Conference on System Sciences, pp. 3005-3014, January 2000