International Journal of Pure and Applied Mathematics
Volume 114 No. 7 2017, 651-661
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue
ijpam.eu

Fuzzy based Data Gathering Scheme for Scalability in WSN

[1]A. Sanjay Gandhi and [2]S. Saira Banu
[1]Dept of ECS, Karpagam University,
Karpagam Academy of Higher Education.
[2]Dept of ECS, Karpagam University,
Karpagam Academy of Higher Education.

Abstract

Wireless Sensor Networks (WSN) consist of sensor node under the control of base station. Data gathering plays a major role in WSN, whereas network lifetime depends on data collection. In dynamic nature, the nodes are get depleted and it leads to low energy because of data lost. To ensure more data collection rate, effective data collection protocol with stable routing is needed. In this research work, Fuzzy based Data Gathering Scheme (FDGS) is developed to increase the data collection rate irrespectively in dynamic environment. It consists of three phases. In the first phase, stable multicast routing is deployed to avoid unstable link failures. In the second phase, cluster head is chosen among the cluster members based on fuzzy decision mechanism. In the third phase, data gathering scheme is deployed to increase the availability of data whenever the node requires. Based on the simulation analysis, the proposed scheme achieves better performance the previous work in terms of packet delivery ratio, data gathering rate, packet loss rate, propagation delay, end to end delay, energy efficiency and communication overhead.

Key Words: WSN, FDGS, data gathering, network lifetime, delay, residual energy and throughput.

1. Introduction

A. Wireless Sensor Networks (WSNs)

Wireless sensor network (WSN) is a group of sensor nodes (SNs) working in uncontrolled areas and organized into cooperative network. It is composed of huge number of sensor nodes which can monitor the environment by collecting, processing as well as transmitting collected data to the remote sink node through direct or multi hop transmission. WSNs have attracted lots of attention in recent years due to their wide applications such as battlefield surveillance, inventory and wildlife monitoring, smart home and healthcare etc. In the previous works , held with the explicit communication

In the previous works [17,18] with the explicit communication approach, fuzzy based optimized routing was used for data gathering based on compression, a communication channel was abstracted as a point-to-point link, which ignored the fact that wireless channels transmitted information by broadcasting it, which made it available to any receiver of the right type. In wireless networks, when one node transmitted data to its destination node, other nodes within the transmission range of the transmitting node could also receive the data and would use it to compress their own information. The proposal of framework for data gathering that exploits the broadcasting characteristic of the wireless medium, thereby achieve greater energy efficiency.

B. Data Gathering Issues in WSN [5]

A dynamic node can act as neighbor node to deliver packets from sensor node to the base station. The data collector collects the packets from sensor node and pass it to the base station. It causes some delay. Energy consumption is a major issue in WSN. Sensor network consists of sensor node that is deployed in network region. The usage of energy plays major role in WSN and it should be utilized effectively.Balancing the data gathering and energy is a major task in WSN. Multiple sensor node takes more data and it can reduce the energy.

C. Challenges of Data Gathering in WSN

Routes are established based on network topology between sensor node and the base station. Without choosing alternative path, the same path can be utilized to pass the data to the base station to save the energy.Many sensor nodes send the data packets to the base station frequently which causes loss of more data, high energy consumption and too delay.

2. Previous Work

Kazuhiko Kinoshita etal. [1] discussed the accurate and adaptive energy estimation method. It was not independent of period length. A data gathering scheme was optimized for the environmental energy-based sensor network .

Sudhakar and Sangeetha [2] initiated the Delay bounded Sink Mobility (DeSM) problem under sensor node allocation to sinks. A polynomial-time optimal algorithm was used for the origin problem. Extended Sink Scheduling Data Routing (E-SSDR) algorithm was also utilized from allocating schedule to the destination node. A mobile sink scheduling scheme was also enhanced to support large size networks.

Roseline & Srinivas [3] proposed secure energy efficient location aware data gathering approach. The properties of node location and energy were used to improve the lifetime of node and network. Elliptic Curve Diffie Hellman Key Exchange was incorporated for securing data gathering between the source and the receiver. The routing was performed based on the neighbor node and the highest energy node selection. The proposed scheme was compared with the existing data gathering scheme without the security measure.

Mehala & Balamurugan [4] presented Clustering-based Data aggregation algorithm for Wireless Sensor Networks using two mobile sink, to increase network lifetime and to reduce the transmission distance between each nodes. Based on this method, the sensed data are collected efficiently which minimizes data latency. Wireless sensor networks are used for application oriented tasks.

Shanmugam et.al [5] introduced an energy efficient mobile element based data gathering using tinybee. It achieved energy in efficient manner and reduced the latency in data gathering. The dispatch of tinybees by mobile element reduced the tour length of the mobile element. This work was determined based on delay and energy. The work achieved in increasing the network lifetime due to less energy utilized than other algorithms.

Anagha & Binu [6] chose the rendezvous point(RPs) from the sensor network according to the weight and other nodes would send their data to this rendezvous nodes in an energy efficient way. the online data collection method was also proposed to reduce the energy consumption during the data collection. Whenever the event was detected by the mobile sink it was not needed to visit every nodes to collect information.

Mingming Lu & Jie Wu [7] modelled the data-gathering problem as an optimization problem. NP-hardness was proposed for both the reverse broadcast tree and the reverse multicast tree problems. Simulation was used to study the effects of different parameters and to compare the performance of various heuristics.

Shilpy Ghai et.al [8] introduced an energy-efficient security protocols to provide strong security and to decrease the energy consumption. The data was compressed and then transmitted it to the other node. With the help of the compression, the energy consumption of the nodes was decreased because fewer amounts of data were spent.

Priyadharshini et.al [9] observed that multiple channel method was helpful in reducing schedule length. The link-based channel assignment schemes had more energy efficient to remove interference. The fundamental limitations due to interference explored techniques reduced the same issue. Once the interference was completely avoided, the achievable schedule length was reduced in the routing tree for aggregated convergecast, and for raw-data convergecast.

Palekar [10] demonstrated a tree based multicast steering convention, MAODV (Multicast Ad hoc On-interest Vector) in lightweight specially appointed systems. The effect of system loaded on MAODV convention, and proposed Multicast Ad hoc On-interest Vector with backup branches to enhance the strength of the MAODV convention by joining focal points of the tree structure and the lattice structure. The shorter tree limbs could enhance as well as build a multicast tree with reinforcement branches.

Liu & Wu [11] proposed the trust management system based on neighbor monitoring. In the trust management system, the trust value was calculated by the neighbor monitoring mechanism, and the direct trust value and the indirect trust value were combined to establish the distributed trust model to detect the malicious nodes. The consistency check algorithm was capable of defending against the attacks on the trust model.

Thenral and Sikamani [12] introduced Angle based Multicast Routing Algorithm (AMRA) for reducing communication overhead and to increase the performance in the wireless mesh networks.

The paper is organized as follows. The Section 1 describes the introduction of WSNs, data gathering and issues. The Section 2 deals with the previous work which is related to the energy efficiency and data gathering algorithms. The Section 3 is devoted for the implementation of proposed data gathering approach. The Section 4 describes the performance analysis and the last section concludes the work.

3. Implementation of Proposed Protocol

In the proposed fuzzy based data gathering scheme, multicast routing is implemented to increase the data availability ratio and provides data integrity in wireless sensor networks. In first phase of the scheme, multicast route is established from Cluster Head (CH) to cluster member or neighbor nodes to provide connectivity. To provide an uninterrupted data communication, stability is integrated in all paths. Cluster heads are also transferring their information via multicast routes.

Stability Based Multicast Routing

Stability means the node or the path can withstand in the presence of dynamic environment, link failures or node failures etc. If the fault tolerant rate of path is high, the stability will also be high.. Path is discovered based on the acknowledgement received from destination node. Initially Cluster Head (CH) floods a Group Join Request (GJR) packet to all cluster members based on the location information.

Stability is calculated based on the signal strength, path capacity and Signal to Noise Ratio [SNR]. Number of repeated transmissions will be limited once the stability is integrated in all the paths. Signal strength is measured in terms of amount of data can be successfully delivered to the nodes. If the delivery rate is more, then signal strength will be high. Path Capacity (P_c) is derived as,

$$P_c = \frac{\Sigma \gamma}{\Sigma \delta}$$

Where γ is number of packets travelling in the path and δ is total number of packets dropped by the paths. If packet dropping is less, path capacity will be more. Stability will be integrated once the high path capacity is arrived. Signal to Noise ratio (SNR) is generally measured based on Bit Error Rate (BER). Both are inversely proportional to each other.

Fuzzy based Cluster Head Election system

Fuzzy method is an effective approach to classify the complex objective in the decision objective domain and to provide optimal solution in the presence of fuzzy logic. Cluster head is chosen based on the following factors i.e Residual energy, number of cluster member nodes connected and stability of nodes. All these factors are included in the simplistic and the objective way. The fuzzy mathematics approach is adopted to obtain inclusive assessment and the attribute weights are found through the entropies based on the positive values of the samples.

In the cluster region, perfect solution and negative solution are estimated. The distance between perfect solution and the attribute samples can be found using weighted Euclidean distance. If any nodes with maximum optimum value, it is chosen as cluster head. The steps for cluster head election using fuzzy method is as follows,

Step 1: $F = \{S_1, S_2 \ldots\ldots S_n\}$ is a distinct set of sensor nodes and $W = \{w_1, w_2 \ldots w_n\}$ is a set of sample attributes for valuation. The viewing matrix V can be calculated as follows,

$$V = \begin{bmatrix} v_{11} & v_{1n} \\ v_{k1} & v_{kn} \end{bmatrix}$$

Where v_{pq} denotes the attribute value of q from node p.Step 2: The computation overhead is reduced based on optimization method. The degree of derivation is given as,

$$z_{pq} = \frac{v_{pq} - \widetilde{v_q}}{\sigma_q}$$

Where $\bar{v_q}$ is mean value of the attributes and $\sigma_q$is standard deviation of attributes.

Step 3: To select the cluster head, the optimum function can be described as follows,

$$\mu = \arg\min_{p \in F} \sum_{x=1}^{N} J_x f_x(p),\ p = 1,2, \ldots ..l$$

Where J_x is weighting factor and it is always greater than 0.

Step 4: Choose the J_{opt} to select cluster head with more residual energy, more number of nodes connected to it.

A. Data Gathering Approach

Data gathering is initiated, once the reference point between the cluster head and the cluster member is identified. It means the centrality degree between two nodes should be less than 45 degree. Cluster member initiates this phase to increase the rate of data availability. The steps for data availability is given as,

Step 1: Choose the reference point between the cluster head and cluster member below 45 degree. Reference point is selected if any event occurs. i.e. transmission or reception.

Step 2: The Cluster member node transmits the Enquiry Packet (EP) to the reference point once the event occurs. Enquiry packet includes the following fields i.e. cluster member ID, cluster member size, distance travelled and hop count.

Step 3: Reference point broadcasts Enquiry Reply Packet (ERP) with zero hop count to cluster member nodes.

Step 4: Once the ERP is received by the cluster member nodes, the hop count will be incremented by 1.

Step 5: If the next hop cluster member node is greater than 1,

Then compare packet arrival time and packet propagation delay Choose the less packet arrival time and propagation delay Start the data transmission

End if,The Cluster head sends the Stable Path Initiation Request (SPIR) to the cluster member. The Cluster member replies via Stable Path Join Reply (SPJR) as next hop node. Cluster members move in the radio range of reference point and it receives gathered packets from the point. Step 6: If an unknown node enters the cluster region, it sends the request packet to CH. CH authenticates it and sends the approval message. CH also measures whether unknown node is inside the coverage region of reference point or not.Step 7: CH calculates the data gathering rate (φ) from the cluster members. It is estimated as follows,

$$\varphi = \frac{\sum_{k=0}^{N} f(w_k)}{\sum w}$$

Where $\sum W$ is total number of packets. $f(w_k)$ means number of packets gathered.Step 8 : CH announces data gathering rate to all the cluster member nodes.

B. Proposed Packet Format

In fig. 1, the packet format of proposed algorithm is shown. The first field source and the destination id occupy 4 bytes. The second field occupies 2 bytes field how much data is collected successfully. In next field, hop count fills 1 byte where it calculates the number of hops from cluster head. Energy efficiency occupies 2 byte field to find how much of energy spent for transmission and reception. Load balancing rate occupies 2 byte field where the data is sent through different paths. Cyclic Redundancy Check (CRC) occupies 4 bytes field for error correction and detection.

The proposed packet format of SDGP is given below;

Source Id and Destination ID	Data gathering rate	Hop Count	Energy Efficiency	Load balancing rate	C RC
2	2	1	2	2	4

Figure 1: Proposed packet format

4. Performance Analysis

Network Simulator (NS2.34) is used to promote the simulate our proposed protocol. Backend language is C++ and Front end language is Tool Command Language (TCL).In the simulation, 120 mobile nodes move in a 1200 meter x 1200 meter square region for 100 seconds simulation time. All nodes have the same transmission range of 260 meters. The simulation settings and parameters are summarized in table 1. Table 2 summarizes the performance metrics comparison of proposed and previous schemes.

Table 1: Simulation Settings and Parameters

No. of Nodes	120
Area Size	1200 X 1200 m^2
Mac	802.11
Radio Range	260m
Simulation Time	100 sec
Traffic Source	CBR
Packet Size	512 bytes
Package rate	5 pkt/s
Protocol	LEACH

A. Performance Metrics

It is evaluated mainly the performance according to the following metrics.**End-to-end delay:** The end-to-end-delay is averaged in over all surviving data packets from the sources to the destinations.**Data Gathering Rate:** It is defined as the making the copies of data items which is shared by several users in a particular point of time.**Communication Overhead:** It is defined as the number of control packets normalized in the network..The simulation results are presented in the next part. The compare of the proposed protocol FDGS with SDGP, STCDG [18] and LEACH in presence of clustering environment.

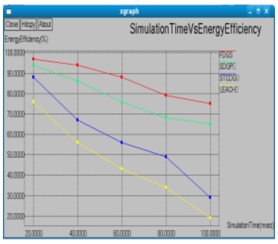

Figure 2: Simulation time Vs energy efficiency

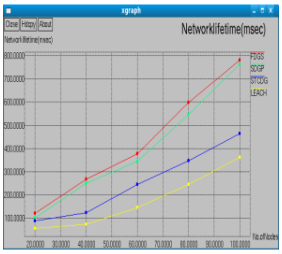

Figure 3: No. of nodes Vs network lifetime

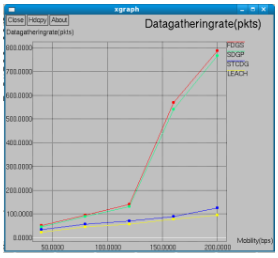

Figure 4: Mobility Vs data gathering rate

Figure 2 shows the results of total energy efficiency for varying the simulation time from 20 to 100 m/sec. From the results, it is understood the scheme FDGS has high energy efficiency than the SDGP, LEACH and STCDG scheme. While adopting the data compression, energy level will be saved. At the same time only active sensor node participates in the data transmission.

Fig. 3, presents the network lifetime comparison for FDGS, SDGP, LEACH, STCDG. It is clearly seen that the number of epochs are consumed by FDGS, is highly compared to SDGP , LEACH and STCDG. Network lifetime is improved based on energy balancing of data gathering nodes and link quality. Fuzzy based decision mechanism is deployed to increase the data gathering.

Fig. 4, presents the comparison of data gathering rate. It is clearly shown that the data gathering rate of FDGS, is higher than SDGP, LEACH and STCDG. The data gathering is improved based on proper maintenance of the cluster region. The Cluster members grasp the data and deliver according to fuzzy routing.

Figure 5 shows the results of Time Vs End to end delay. From the results, it is clearly shown that FDGS scheme has slightly lower delay than SDGP, STCDG and LEACH scheme because of stable data gathering routes.

Fig. 6, presents the comparison of overhead while varying the mobility from 20 to 100 bps. It is clearly shown that the delay of FDGS is lower than SDGP, STCDG and LEACH. Due to the implementation of data gathering mechanism, the packet replay is avoided to reduce the packet overhead. Hence the overall network communication overhead will get reduced.

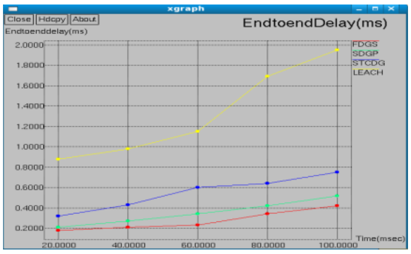

Figure 5: time Vs end to end delay

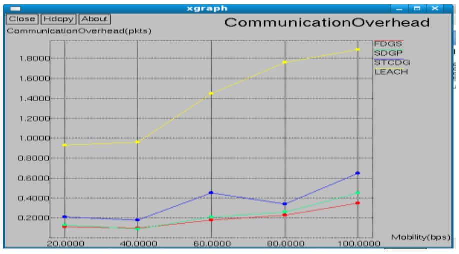

Figure 6: mobility Vs communication overhead

5. Conclusion

In WSNs, the stable route can be determined based on the link failures between neighbor nodes. To increase the network lifetime, many sensor nodes should be deployed to increase data gathering rate. To perform data gathering, the fuzzy decision routing is required to achieve data availability. In this paper, balancing between the energy consumption and the data gathering is achieved through fuzzy based data gathering route which attains energy model and attains high network life time and throughput to the sensor nodes. it has been demonstrated the data gathering estimation of each node. By simulation result it has been shown that FDGS has achieved high data gathering rate, high network lifetime, high energy efficiency while attaining low delay and low overhead than the existing schemes LEACH and STCDG and the previous work SDGP while varying the number of nodes, simulation time, time and mobility.

References

[1] Kazuhiko Kinoshita, Takahisa Okazaki, Hideki Tode and Koso Murakami, A Data Gathering Scheme for Environmental Energy-based Wireless Sensor Networks, IEEE Communication Society (2008), 719-723.

[2] Sudhakar B., Sangeetha K., Multi Sink based Data Collection Scheme for Wireless Sensor Networks, International Journal of Innovative Research in Computer and Communication Engineering 2(1) (2014), 1139-1146.

[3] Roseline Juliana M., Srinivasan S., SELADG: Secure Energy Efficient Location Aware Data Gathering Approach For Wireless Sensor Networks, International Journal on Smart Sensing and Intelligent systems 8(3) (2015), 1748-1767.

[4] Mehala R., Balamurugan A., An Efficient Data Aggregation Scheme and Cluster Optimization in Wireless Sensor Networks, International Journal of Innovative Research in Computer and Communication Engineering 3(3) (2015), 1706-1712.

[5] Shanmugam P., Jayanthi S., Raja J., Nusrath R., Energy Efficient Mobile Element Based Data Gathering in Wireless Sensor Networks, International Journal of Innovative Research in

Electrical, Electronics, Instrumentation and Control Engineering 3(11) (2015).

[6] Anagha K., Binu G.S., An on demand Data Collection Scheme for Wireless Sensor Network Based on Rendezvous Points, International Research Journal of Engineering and Technology, 2(4) (2015), 78-82.

[7] Mingming Lu, Jie Wu, Utility-Based Data-Gathering in Wireless Sensor Networks with Unstable Links, NSF grants ANI (2007).

[8] Shilpy Ghai, Katiyar V.K., Energy Efficient Data Transmission Schemes in Wireless Sensor Networks, International Research Journal of Innovative Engineering 1(4) (2015), 1-4.

[9] Priyadharshini K., Bhuvenswari S., Arun R., Cholaraja K., Fast Data Collection Using Two-Layer Multicast Communication Protocol in Wireless Sensor Networks, International Journal of Innovative Research in Computer and Communication Engineering 3(3) (2015), 2510-2516.

[10] Vikas R. Palekar, Tree structure based Optimized Multicast Routing Algorithm in MANET, International Research Journal of Engineering and Technology 2(2) (2015), 471-479.

[11] Bei Liu, Yuanming Wu, A Secure and Energy-Balanced Routing Scheme for Mobile Wireless Sensor Network, Wireless Sensor Network 7 (2015), 137-148.

[12] Thenral B., Thirunadana Sikamani K., AMRA: Angle based Multicast Routing Algorithm for Wireless Mesh Networks, Indian Journal of Science and Technology 8(13), 1-8.