Context Aware Data Aggregation in Distributed Sensor Networks

Lokesh. B. Bhajantri

Department of Information Science and Engineering,
Basaveshwar Engineering College, Bagalkot, Karnataka,
lokeshcse@yahoo.co.in

Nalini. N

Department of Computer Science and Engineering,
Nitte Meenakashi Institute of Technological, Bangalore,
nalinaniranjan@hotmail.com

Abstract – Distributed Sensor Networks (DSNs) can provide access to information anytime, anywhere by collecting, processing, analyzing, and disseminating data. Thus the network actively participates in creating a smart environment. DSNs promise to revolutionize sensing in a wide range of application domains. An equipped with appropriate distributed sensors in DSN that can enable the detection of objects, temperature, and identifying the location of fire attack in the forest. Sensor nodes deployed randomly in the forest environment to gather the context information such as air pressure, temperature context, objects aware, location of fire, fire condition (emergency level or non emergency level), and energy aware of each node.

In this work, based on the context, sensor nodes configure, sense, and send the aggregated context data to base station (sink node). In this work, identifies four different contexts in DSN by considering the scenario of forest environment such as: temperature context, air pressure context, energy aware context, and object aware context. The two levels of data aggregation processes (Cluster Head node and Sink node levels) are considered for minimizing the redundant data for transmission. Finally, we evaluated the performance parameters for the proposed scheme.

Keywords-: Distributed Sensor Networks (DSNs), Context Aware, Node Knowledge Base (NKB), Cluster Head (CH).

I. INTRODUCTION

A Distributed Sensor Network (DSN) is a collection of a large number of heterogeneous intelligent sensors distributed logically, spatially, or geographically over an environment and connected by a high speed network. This is because of their reliability, accuracy, flexibility, cost effectiveness, and ease of deployment of nodes in DSN. Smart sensors can offer vigilant surveillance and can detect and collect data concerning any sign of machines failure, earthquakes, fire attacks, and floods and even a terrorist attack.

Sensor nodes periodically sense the data from the surrounding environment, process, and transmit it to the base station or sink node. The frequently sensing of data and the number of sensors periodically reports the data that depends on the particular applications. Data aggregation involves systematically collecting the sensed data from multiple sensors, aggregating to suppress the redundant data, and transmitting the aggregated data to the sink node for further processing [2] [3]. The main objective of the data aggregation is to eliminate the redundant data, transmission to prolong or conserve the energy of the sensor nodes of DSN. Hence, it is needed to combine the sensed data into high quality information and this is accomplished through data aggregation [4] [5] [6].

The context environment leads to DSN devices or nodes in contexts, i.e., the ability of device or program to sense, react or adapt to its environment to use. The optimal heuristic techniques are used to solve complex real time problems such as context aware devices, context discovery, sensing, extraction and manipulation, which have a high degree of mechanism for providing flexible services as well as quality of services. This information can be utilized to create smarter environment to achieve the challenges like monitoring and collecting the data, assessing and evaluating the information, formulating meaningful information, and performing decision making [7].

The objective of this work is to design and analyze the context aware data aggregation in DSNs for handling trade-offs i.e. trade-offs between the different objectives such as energy consumption, aggregation time, routing overhead, and throughput, improving quality of service of the data aggregation protocols in terms of bandwidth and end to end delay.

The rest of the paper is organized as follows: Section II presents an overview of related work. The proposed work is discussed in Section III. Simulation and results analysis are presented in section IV and finally we conclude the proposed work in section V.

II. RELATED WORK

Some of the related works are as follows: The work given in [8] presents a grid based data aggregation (GBDA) for WSNs. In this paper, the grid is divided into number of grid cells in the WSN environment for minimizing amount of data transmitted to the base station. In this work, data is aggregated at each grid cell heads and resultant data is send to the base station. The proposed scheme also performs data aggregation for each cell heads in the chain except both end nodes and reduces data transmissions to the base station.

The work given in [9] discusses context aware routing protocols used in DSN. In which routing techniques are used to implement in different applications. The paper [10] presents a routing in WSN. In this work, wireless sensor networks are formed by small sensor nodes communicating over wireless links without using a fixed network infrastructure. Different geographic routing strategies are described as well as beacon less routing techniques.

The work given in [11] presents a generic framework for context-aware routing and its implementation in WSNs including the message flow as well as the route decision method. The framework is not limited to a single application domain, but it is a generalized approach that can be adapted and specialized for various application domains. According to [10], DSN are designed for many monitoring and surveillance tasks. It discusses about routing framework in WSN to aware of environmental monitoring.

The work given in [12] presents a data aggregated maximum lifetime routing scheme for WSNs. In which address the problem of jointly optimizing data aggregation and routing. A recursive smoothing method is adopted to overcome the non-differentiability of the objective function. Derive the necessary and sufficient conditions for achieving the optimality of the optimization problem and design a distributed gradient algorithm accordingly.

The work given in [13] describes that, in WSN efficiently disseminating data from a dynamic source to multiple mobile sinks is important for applications such as mobile target detection and tracking. Some of the related works are given in [14] [15] [16] [17] [18] [19].

The contributions of the proposed scheme are as follows: (1) Context information are gathered from different context manager nodes (sensor nodes) as well as generated at sink node. (2) Forest scenario or environment is monitored using proposed method for identifying the location of fire attack. (3) Used context manager nodes to perform several tasks that aid the information processing, data gathering, data aggregation, and data dissemination etc., in asynchronous fashion. (4) Based on context, low or high fire detection in employed. (5) Data gathering and dissemination from active sensor nodes. (6) Two level of data aggregation process to eliminate the redundant data more efficiently: local aggregation (context manager nodes) and master level aggregation (sink node level). (7) In which, context is in emergency level, data is disseminated to target area, and (8) Finally, simulated the proposed scheme to test the operation scheme in terms of performance parameters for the proposed model.

III. PROPOSED WORK

The proposed context aware data aggregation (CADA) inspired routing protocol which shows enhanced performance in terms of energy utilization, routing overhead, throughput, and time taken for data aggregation, the schemes by transmitting the aggregated data from the sensor nodes to the base node or sink node. In the proposed work, context aware data aggregation scheme is evaluated by considering forest scenario. The numbers of sensor nodes are deployed randomly in heterogeneous DSN environment. The network is divided in to number of clusters namely, $C_N = C_1, C_2, C_3, \ldots C_N$. In which number of context types were considered namely temperature aware, energy aware, air pressure aware, and object ware contexts. In each clusters, cluster head (CH) is elected based on highest energy among the nodes. CH is used to monitor the status of the clusters of the DSN environment in terms of temperature, energy of each sensor nodes, air pressure in the cluster, and objects in the clusters. CH nodes are used to collect relative information from the respective clusters from its members for efficient data aggregation process. This section presents network environment and functioning scheme for data aggregation.

The Fig. 1 shows the network environment for the proposed work. The proposed system model consists of set distributed sensor nodes that are deployed randomly with diversified sensing competence and sink node in DSN. The network is divided into number of clusters. Each cluster comprises of set of contexts. All the CHs collect the following information:

- To collect the temperature of cluster by using its cluster member nodes. CH aggregates the data.
- To collect energy information of each node of its cluster. CH maintains the information of energy level of each node.
- To collect the information of objects from the cluster in the form Boolean attributes. There may or may not be objects in the cluster. i.e., TRUE or FALSE.
- To collect the air pressure information in the cluster by the sensor nodes. Air pressure differs in every fraction of seconds, this random information is maintained by CH.

The operation sequence of proposed system model is as follows: (1) Placing of CH node in each cluster for monitoring, (2) the two levels of aggregations take place in this work namely one is aggregation at each CH (local aggregate) and another is aggregation at the sink node level (master aggregate), (3) sink node takes action upon receiving the information from CH manager nodes for data dissemination. (4) Master aggregator analyzes the data, and based on the values, it decides the context whether it is in emergency level or in non-emergency level of context by the sink node, and (5) sink node takes the action upon receiving the information from the sensor nodes in DSN.

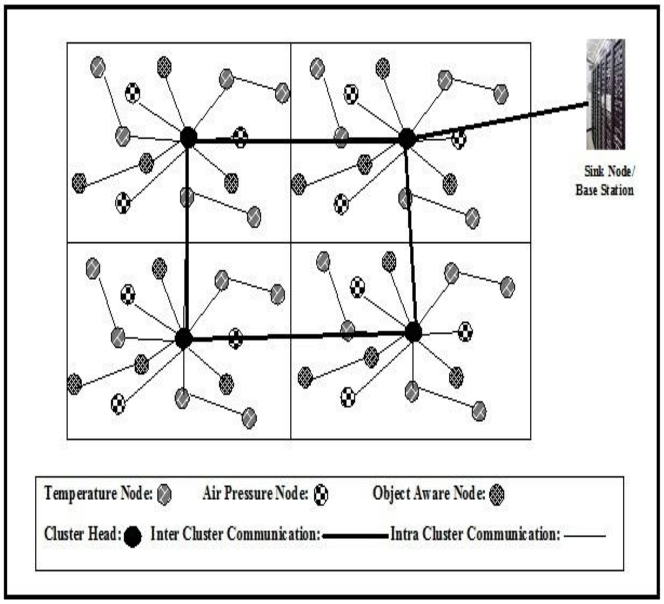

Fig. 1. Network Environment.

The CHs and Sink node maintains the Node Knowledge Base (NKB). This knowledge base is read and updated by the sensor nodes or CHs. NKB comprises of Sensor node id, active mode/sleep mode, emergency level, previously aggregated data, set of critical data, bandwidth information, number of cluster member node, the number of CHs, number of clusters, sink node information, and location of the node. The sensor node senses the data and updates the NKB. The detail view of NKB is shown in table 1.

Table 1: Node Knowledge Base

C_ID	CH_ID	S_{id}	Status	E_S	C_B	Contexts	Emergency Level
C_1	CH_1	S_1, S_3, S_5, S_7, S_{16}, S_{18}, S_{19} . .	Active Sleep	High Medium Low	High Average Low	S_N = Temperature S_N = Object Aware S_N = Air Pressure	High Medium Low
C_2	CH_2	S_2, S_4, S_6, S_8, S_{10}, S_{11}, S_{13} . .	Active Sleep	High Medium Low	High Average Low	S_N = Temperature S_N = Object Aware S_N = Air Pressure	High Medium Low
. . . . . .	. . . . . .	. . . . . .	. . . . . .	. . . . . .	. . . . . .	. . . . . .	. . . . . .
C_N	CH_N	S_N	------	------	------	------	------

C_ID = Number of Clusters, CH_ID = Number of Cluster Heads, S_{id} = Number of Sensors, E_S = Energy of Sensor Nodes, C_B = Cluster Bandwidth

A. Functioning Scheme: Proposed Algorithm

Nomenclature:

{Area = 5000* 5000 meter, S_N = 100 to 500 nodes, C_N = number of clusters, S_{id} = Sensor ID, Es = energy of sensor node, T_R = Transmission Range, N_b = Bandwidth of the Network, Temperature (T) = min-20°, max-100°c, Air-Pressure (AP) = Nm^{-2}, Transmission Range (T_R) = 200m, number of sink node (N_S) = 1, H = High, M = Medium, L = Low, σ = Air Density, p = Absolute Pressure, RSpecific = The specific constant, and GPS = Global Positioning System}

Begin

1. Deploy number of nodes in an environment randomly;
2. Configure the DSN with S_{id}, E_S, $T_{R,}$ GPS, and $N_{b;}$
3. The DSN is categorized into number of clusters;
4. Select the CHs Nodes in each clusers;
5. Proposed protocol CADA is used in the data aggregation process is as follows: Let 'β' is the data aggregation factor at CH nodes in each clusters, 'δ' is the resultant aggregation factor in each cluster is as follows:

$$\delta 1 = \sum_{c=1}^{n} \beta(\mathbf{CH}(\mathbf{S1\ U\ S2\ U \ldots\ldots U\ SN})) \quad \textbf{for Temperature}$$

$$\delta 2 = \sum_{c=1}^{n} \beta(\mathbf{CH}(\mathbf{S1\ U\ S2\ U \ldots\ldots U\ SN}))$$

for Energy

$$\delta 3 = \sum_{c=1}^{n} \beta(\mathbf{CH}(\mathbf{S1\ U\ S2\ U \ldots\ldots U\ SN})) \quad \textbf{for object aware data}$$

$$\delta 4 = \sum_{c=1}^{n} \beta(\mathbf{CH}(\mathbf{S1\ U\ S2\ U \ldots\ldots U\ SN}))$$

for Air Pressure Aware

The AP information is handled by CH node, which is calculated by using following equation.

$$\sigma = \frac{p}{(RSpecific \times T)}$$

Therefore,

AP α T

6. CHs are send the aggregated data (Temperature (δ1), Energy (δ2), Object Aware (δ3), and Air pressure (δ4)) to sink node;
7. The MA at the sink node, 'μ' is the resultant MA factor, 'β' is the aggregation factor for the cluster data is computed by using the following equation. Therefore, sink node decides the emergency level or non emergency level of context in DSN.

$$\mu = \sum_{c=1}^{n} \beta(\mathbf{C1\ U\ C2\ U \ldots\ldots U\ CN})$$

Then,

$$\mu = \begin{cases} if(T \geq 100 \ \&\& \ AP = H \ \&\& \ Object = True \\ if(T \leq 100 \ \&\& \ AP = M \ \&\& \ Object = True \\ if(T \geq 30 \ \&\& \ AP = L \ \&\& \ Object = True \end{cases}$$

.

8. The sink node decides the level of context based on the value of 'μ', whether high emergence level cluster or medium emergence level cluster or non-emergence level cluster in the network.
9. If the sink decides the emergency levels, then sink node broadcast the data to the targeted clusters.

End

IV. SIMULATION

The simulation of the proposed scheme is simulated by using 'C' programming language. The proposed scheme has been simulated in various network scenarios. Simulations are carried out extensively with random number for 1000 iterations. This section presents the simulation model, simulation procedure, performance parameters, results and discussions.

A. Simulation Model

The simulation model consists of 'N' number of sensor nodes deployed randomly in a distributed environment. Simulation is done for 'S_N' (S_N=100) sensor nodes used to measure the performance parameters such as energy utilization, redundancy ratio, time taken for data aggregation, and routing over head.

B. Simulation Procedure

Table 2 presents the simulation parameters considered for analyzing the scheme.

Table 2: Simulation Inputs

Parameters	***Notations***	***Values***
Length	l	5000 meters
Breadth	b	5000 meters
Number of Sensor Nodes	S_N	100 – 500 nodes
Sensor node Transmission range	T_R	200 meters
Initial Energy of sensor nodes	I_E	20 Joules
Size of packets	S_P	32, 64, 128, 256, 512, 1024 Kbytes
Energy required for sensing of data	E_S	50nJ/Bit
Energy required for transmission of data	E_T	50nJ/Bit
Threshold level energy	THLE	0.05J.
Number of Sink Node	S_N	1

Simulation procedure involves following steps:

Begin

1) Deploy the number of sensor nodes in DSN randomly;
2) Initialize the sensor node properties;
3) Initialize CH node in each clusters;
4) Apply the data aggregation algorithm;
5) Compute the performance parameters;

End

C. Performance Parameters

The following parameters are used to measure the performance of the proposed scheme:

1) *Energy Utilization:* As the number of rounds increases, there is decrease in the energy of sensor nodes in DSN.
2) *Redundancy Ratio:* It is the ratio of size of resultant redundancy data (after aggregation) to the total size of the data and is defined as follows:

   Redundancy Ratio = Size of resultant redundancy data / Total size of the data

3) T*ime Taken for Data Aggregation:* As the number of nodes increase, aggregation time also increases in DSN. It is measured in terms of 'mseconds'.
4) *Routing Overhead:* As the number of nodes increase, the routing overhead also increases for dissemination of data in DSN. It is measured in terms of percentage (%).

D. Results and Discussions

Fig. 2 presents the energy utilization for sensor nodes in each round. As the number of rounds increases, the energy of each node decreases in DSN. The maximum energy of each node is considered as 20J in DSN. If the energy is less than the threshold level energy, then node is sent to sleep mode in the network.

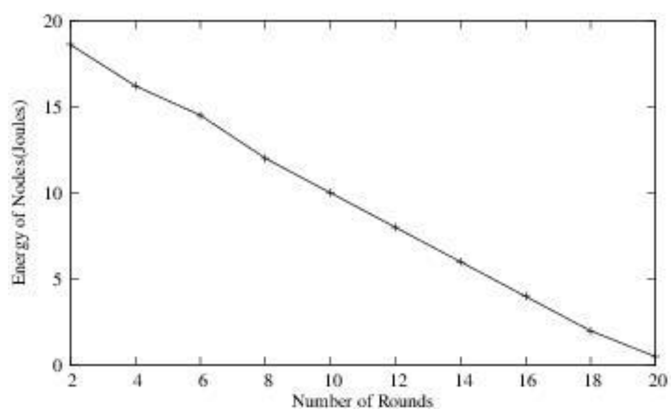

Fig. 2. Energy Consumption Vs. Number of Rounds.

Fig. 3 presents the redundancy ratio for sensor nodes in DSN. As the number of sensor node increases in DSN environment, the redundancy ratio increases. As compared to non-context aware system, the proposed scheme performs better in redundancy ratio because numbers of sensor nodes are configured with temperature, air pressure, object aware and energy aware contexts in DSN.

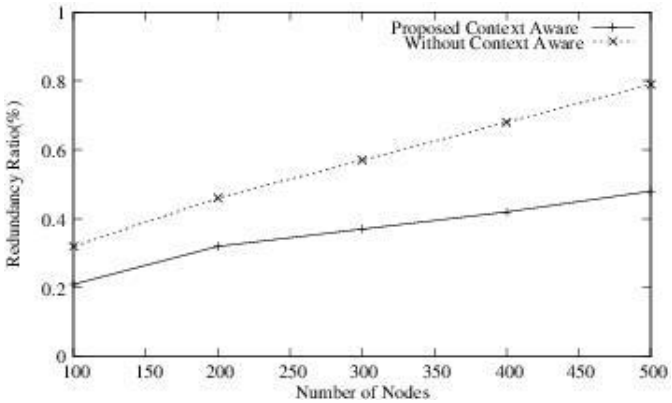

Fig. 3. Redundancy Ratio Vs. Number of Nodes.

Time taken for aggregation among with number of nodes is shown in Fig. 4. As the number of nodes increase, the time for data aggregation also increases in DSN. In this work, sensor nodes are configured with different contexts. Hence context used reduces the time for data aggregation since the data is gathered according to the requirements in the network although the redundancy ratio is less.

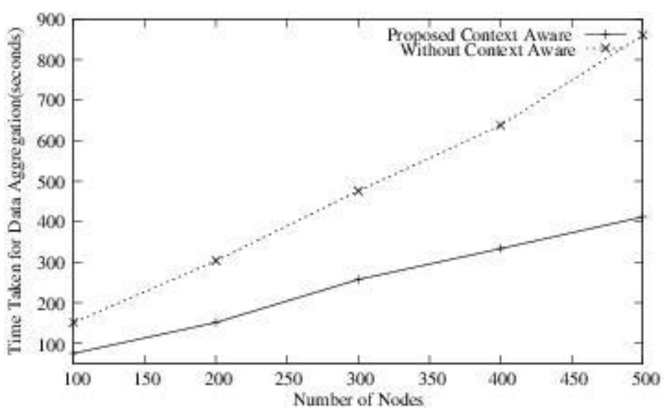

Fig. 4. Time Taken for Data Aggregation Vs. Number of Nodes.

Fig. 5 shows the routing overhead with number of nodes. As the number of sensor nodes increase, routing overhead increases in dissemination of data because of increase in the number of nodes that creates the complexity in the network. For network with less number of nodes, the dissemination rate is high. The proposed scheme uses less number of routing overhead as compared with non context aware system.

Fig. 6 shows the throughput of the network. As the number of sensor nodes increase, there is gradual decrease in the throughput of the data disseminated by the sink node because of congestion in the network. The proposed scheme performs better throughput, as compared to non context aware system since all the sensor nodes are configured with different types of contexts in DSN.

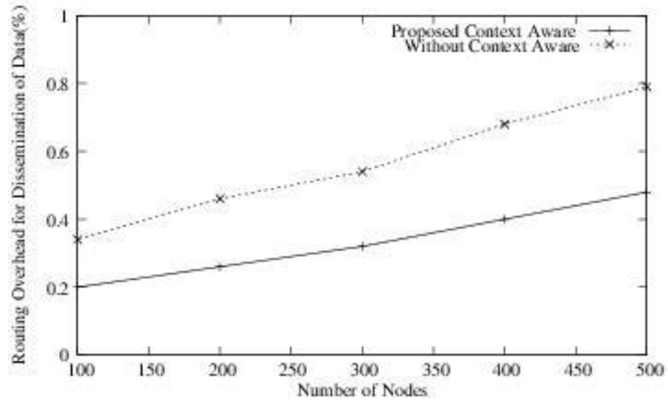

Fig. 5. Routing Overhead Vs. Number of Nodes.

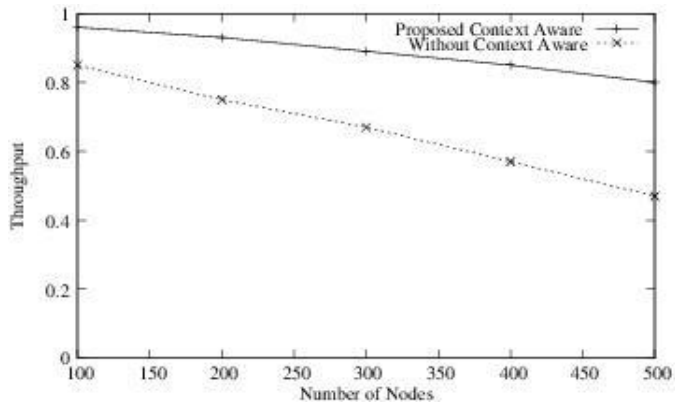

Fig. 6. Throughput Vs. Number of Nodes.

V. CONCLUSION

In this work, the proposed a CADA protocol is used to aggregate the context data at two levels and then sink node disseminates the data to the targeted node. The proposed model considers the context information such as temperature, air-pressure, energy, and object aware contexts. The proposed model is efficient for identifying the location of fire attack in the forest. The proposed model is evaluated by comparing with performance parameters like aggregation time, redundancy, routing overhead, throughput, and energy utilization. A simulation result shows that context information of our proposed system is more efficient than results without context.

REFERENCES

[1] Shivakumar. S, Iyengar. S., S, "Taxonomy of Distributed Sensors Networks". pp. 29 - 43, 2007. Available: http://systems.ihp-microelectronics.com/uploads/downloads/DS1-2007- 06-Taxonomy.pdf

[2] Ihsan. A, Rahim. K, Mehdi. H, Mustaq. A, "Energy Aware Routing and Data Aggregation in Wireless Sensor Networks", In Proceedings of 2nd International Conference on Machine Learning and Computer Science, pp. 27 - 30, 2013.

[3] Bhaskar. K, Deborah. E, Stephen. W, "The Impact of Data Aggregation in Wireless Sensor Networks", In Proceedings of International conference on Distributed Computing Systems, pp. 575 - 578, 2002.

[4] Ramesh. R, Pramod. K., V, "Data Aggregation Techniques in Sensor Networks: A Survey", IEEE Communications Surveys and Tutorials, Vol. 8, No. 4, pp. 48 - 63, 2006.

[5] Yee. W., L, Marimuthu. P, Raphael. C., P, "Secure Data Aggregation in Wireless Sensor Networks", Wireless Sensor Networks, Computer Communications and Networks, Springer-Verlag, Chapter 21, pp. 533 - 560, 2009.

[6] Nidhi. G, Sanjeev. S, Renu. V, "Data Dissemination in Wireless Sensor Networks: A State-of-the Art Survey", International Journal of Computer Networks , Vol. 4, No. 1, pp.22 - 34, 2012.

[7] Chong. S., K, Krishnaswamy, S, Loke. S., W, "A Context Aware Approach to Conserving Energy in Wireless Sensor Networks", In Proceedings of 3rd International Conference on Pervasive Computing and Communications, pp. 401 - 405, 2005.

[8] Neng-Chung. W, Yung-Kuei. C, Chih-Hung. H, Young-Long. C, "Grid-Based Data Aggregation for Wireless Sensor Networks", Journal of Advances in Computer Networks, Vol. 1, No. 4, pp.329 - 333, 2013

[9] Arturas. L, "Context Routing in Wireless Sensor Networks", it Technical Report, pp. 1 - 29, 2006. Available: www.lukosius.netifiles/arturas-CARreporLvl.O-final.pdf

[10] Akkaya. K, Younis. M, "A Survey on Routing Protocols for Wireless Sensor Networks", Elsevier Journal of Adhoc Networks, Vol. 3, No. 3, pp. 325 - 349, 2005.

[11] Wenning B., L, Timm-Giel. A, Gorg. C, "A Generic Framework for Context Aware Routing and its Implementation in Wireless Sensor Networks", In Proceedings of International Conference on Technology and its Applications, pp. 53 - 58, 2009.

[12] Cunqing. H, Tak-Shing. P., Y, "Data Aggregated Maximum Lifetime Routing for Wireless Sensor Networks", Elsevier Journal on Adhoc Networks, Vol. 6, No. 3, pp. 380 - 392, 2008.

[13] Ugur. C, Andrew. F, Sun. Y, "Power Efficient Data Dissemination in Wireless Sensor Networks", In Proceedings of 3rd ACM International Workshop on Data Engineering for Wireless and Mobile Access, pp. 1 - 8, 2003.

[14] Bernd. L., W, Dirk. P, Andrew. T, Carmelita. G, "Environmental Monitoring Aware Routing in Wireless Sensor Networks", In Proceedings of International Conference on Mobile and Wireless Communications Networks and Personal Wireless Communications, pp. 5 - 16, 2008.

[15] Wensheng. Z, Guohong. C, Tom. L., P, "Dynamic Proxy Tree Based Data Dissemination Schemes for Wireless Sensor Networks", Journal on Wireless Networks, Vol. 13, No. 5, pp. 583 - 595, 2007.

[16] Jeongho. S, Tae. Y., B, "A Routing Scheme with Limited Flooding for Wireless Sensor Networks", International Journal of Future Generation Communication and Networking, Vol. 3, No. 3, pp. 19 - 31, 2010.

[17] Arijit. B, "Exploration and Data Aggregation in Sensor Networks", Technical Report, pp. 1- 80, 2007. Available: www.facweb.iitkgp.emet.in/pabitraifacad/06CS601Ot.pdf

[18] Chamran. A, Javad. A., T, "A New Approach to the Data Aggregation in Wireless Sensor Networks", International Journal of Computer Science, Vol. 9, No 3, pp. 328 - 335, 2012.

[19] Lokesh. B., B, Nalini. N, "Optimization of Routing in Distributed Sensor Networks using Heuristic Technique", International Journal of Computer Network and Information Security, Vol. 5, No. 11, pp. 42 - 50, 2013.

AUTHORS PROFILE

Lokesh B. Bhajantri received P.h.D (Computer Science and Engineering) degree from Visveswaraiah Technological University (VTU), Belgaum, Karnataka, 2015. He is presently working as a Assistant Professor in the Department of Information Science and Engineering, Basaveshwar Engineering college, Bagalkot, India, He has experience of around 12 years in teaching and research. His areas of interest include Distributed Sensor Networks (DSNs), Cognitive Internet of Things (CIoT), e-commerce, u-commerce, mobile computing and communications, networking protocols, genetic algorithms, applications of agents and real time systems. He has given invited lectures in AICTE sponsored workshops and seminars. He has published chapter in Handbook of Research on Telecommunications Planning and Management for Business, 12 referred international conferences papers and 10 referred international Journals. He is a reviewer of some journals and conferences. He is a member of Board of Studies (BOS) in the Department of Information Science and Engineering, Basaveshwar Engineering College, Bagalkot, Karnataka, India. He is a member of International Association of Computer Science and Information Technology (IACSIT).

Nalini. N received Ph.D from Visvesvaraya Technological niversity, Balgaum, Karnataka, India. She is currently working as Professor a nd Head of Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology (NMIT), Bangalore, Karnataka, India. She has experience of around 22 years in teaching and research. She is involved in research of wireless and Distributed Sensor Networks (DSNs), Cloud computing, Cryptography & Network Security, Genetic Algorithms, and Heuristic Algorithms in Secure Networks. She is an associate editor of Research Journal of Information Technology, Maxwell Scientific Organization. She has many given invited lectures and has conducted several workshops/seminars/conferences. She has published 30 referred Journals and about 50 referred conferences papers. She is a reviewer of many journals and conferences.