ENERGY AWARE MODEL FOR SENSOR NETWORK: A NATURE INSPIRED ALGORITHM APPROACH

Doreswamy and Srinivas Narasegouda

Department of Computer Science, Mangalore University, Mangalore, India

ABSTRACT

In this paper we are proposing to develop energy aware model for sensor network. In our approach, first we used DBSCAN clustering technique to exploit the spatiotemporal correlation among the sensors, then we identified subset of sensors called representative sensors which represent the entire network state. And finally we used nature inspired algorithms such as Ant Colony Optimization, Bees Colony Optimization, and Simulated Annealing to find the optimal transmission path for data transmission. We have conducted our experiment on publicly available Intel Berkeley Research Lab dataset and the experimental results shows that consumption of energy can be reduced.

KEYWORDS

Ant Colony Optimization, Bees Colony Optimization, Data Mining, DBSCAN, Sensor Network, Simulated Annealing.

1. INTRODUCTION

Sensor networks are deployed in the real world to acquire measurements at distinct points. The acquisition measurement by a sensor node is characterized specifically by two dimensions namely time and location. Recent developments in technologies have developed smart devices characterized by sensing, computational, and communication capabilities. These smart devices have been exploited in various applications such as traffic monitoring [1], surveillance [2], habitat monitoring [3], fire detection, pollution monitoring, and environmental monitoring etc for monitoring the given area continuously. To effectively accomplish this task, the acquisition of data from all the sensor nodes describing the state of the monitored environment is performed [1], [4-5]. Since the main cause for consumption of energy of sensor node are data acquisition and communication [6], these approaches are considered as highly energy consuming. Hence, smart and intelligent techniques are required to reduce the consumption of energy in sensor nodes by focusing the development of energy aware models during the data acquisition.

This paper describes the framework for the development of energy aware model for sensor network by minimizing the communication cost. Density based clustering technique such as Density Based Spatial Clustering of Applications with Noise (DBSCAN) is used to exploit the temporal correlation among the sensor data and spatial correlation among the sensor nodes. Among the correlated sensor data, a subset of sensor nodes is identified. Instead of querying all sensor nodes, only subset of sensor nodes is queried to reduce the amount of data to be acquired and transferred. To minimize the communication cost, travelling salesmen problem solver is used [7]. Experiments were performed on the publicly available Intel Berkeley Research lab dataset [8] to show the effectiveness of the proposed model.

DOI : 10.5121/ijdms.2014.6403

The rest of the paper is planned as follows. Section 2 gives the literature survey; Section 3 describes the proposed model. Section 4 explains the experimental results and conclusion is given in Section 5.

2. LITERATURE SURVEY

Many researchers have devoted their precious time to develop the techniques for the better power management of sensor network. These techniques can be classified into three classes: (i) Reducing the number of transmission, (ii) reducing the number of sensor nodes required to answer the queries, (iii) identification of correlated sensors by exploiting the clustering techniques.

To detain the correlations and statistical relationships among the collected sensor attributes, a probabilistic approach was proposed in [9]. In this, a pull based approach is used to reduce the communication cost. But, in case of anomaly events, this approach does not respond in time. Based on the spatial and temporal correlations, a probabilistic model is proposed in [10]. In this approach communication is reduced by querying the network is only when sensors sense the values outside the error bound. By exploiting the forecasting the time series, a Probabilistic Adaptable Query (PAQ) system was proposed in [11]. In this model each node consists of a local probabilistic model whose parameters are notified to the base station. The base station uses all the local probabilistic models to predict the readings of each sensor nodes. When a sensor sense the reading which is not expected, it relearning phase is executed to build new model and the new model's information is sent to the base station.

To reduce the number of sensor nodes to be queried by selecting a subset of sensor nodes which represents the entire network state was first proposed by [12]. Sensor nodes exchange messages between their spatially correlated neighbours to select the representative sensors of their environment. This technique has been improvised in [13] by exploiting the temporal correlation among the sensor readings. But these techniques fail to identify the correlation among the sensors deployed far away.

Clustering techniques such as PREMON [14], LEACH [15], and CAG [16] have been exploited for the identification of subset of sensor nodes. In PREMON, MPEG compression algorithms are exploited in the cluster head node to predict the readings. Energy consumption is reduced by not transmitting the predicted data. In LEACH, randomized rotation of cluster heads ensures the energy overhead among the sensor nodes in the network. Given a spatial correlation threshold, CAG discovers the cluster of nodes by analyzing the sensed data. As long as the readings are within the threshold, only one reading per cluster is transmitted resulting in reduction in the amount of data acquiring and approximation of aggregation of results within in the threshold. However, these clustering techniques have their own limitations such as prior knowledge of clustering topology is required in PREMON, correlation among far away sensors cannot be detected in LEACH, and only one reading per cluster is transmitted in CAG.

3. PROPOSED MODEL

The proposed model consists of three phases. In the first phase, clustering technique is used to exploit the spatial and temporal correlation among the sensors. In the second phase, subset of sensor nodes is identified for each cluster. And in third phase, optimal transmission path is calculated for the data transmission using nature inspired algorithms such as Ant Colony Optimization (ACO), Bees Colony Optimization (BCO), and Simulated Annealing (SA).

3.1. Spatiotemporal relational model

Spatiotemporal relation model is one which identifies the correlation among the sensors in both spatial and temporal dimensions. For this it exploits the clustering technique. In particularly, Density Based Spatial Clustering of Applications with Noise (DBSCAN). DBSCAN is selected over the partition based clustering techniques because DBSCAN can identify the spherical and non-spherical shaped clusters, it is less sensitive in presence of outliers, and it does not require to specify the number of clusters to be formed in advance. Once the DBSCAN is applied, outliers are identified and removed if there is any. The detailed description of DBSCAN algorithm is given in [17].

3.2. Identification of Representative Sensor Nodes

Identification of subset of sensor nodes plays a very important role in reducing the consumption of energy in sensor network. In this paper we have used the same technique as used in [18-19] for the identification of representative sensor nodes. Once we identify the subset of sensor nodes called representative sensors, instead of acquiring and transmitting data from all nodes only representative sensor nodes are used for acquiring and transmitting the data which results in reduction in the consumption of energy.

3.3. Optimal Transmission Path for Data Transmission

In a large sensor network data is transmitted to the base station through multiple hops. If the signal strength between any two sensor nodes is weak then packet may get lost resulting in retransmission of data which further increases the energy consumption. Hence, it is very important to identify the path whose signal strength is more compared to the other paths. But finding the optimal transmission path among all possible paths is a combinatorial problem which is very time consuming. Nature inspired algorithms such as Ant Colony Optimization (ACO), Bees Colony Optimization, and Simulated Annealing (SA) have been used in the past for solving combinatorial problems. In this paper we also used these nature inspired algorithms to identify the optimal transmission path for data transmission. The detailed description of nature inspired algorithms can be found in [20]

3.3.1.Ant Colony Optimization

Ant Colony Optimization technique was introduced by [21-22]. ACO is inspired by the behaviour of ants in the real world. In the real world, ant looks for the food and when food is found they return to their home. While returning to their home they lay the pheromone along the path. As the path is used more, more pheromone is collected along the path. And soon other ants also start following the same path. Soon or later all the ants will be travelling between the food source and their home on an optimal path.

In ACO, the pheromone and problem specific heuristic information is used to construct the probabilistic solution. For a given component *i*, the probabilistic solution is given by

$$P_{ij} \leftarrow \frac{\tau_{i,j}^{\alpha} \times \eta_{i,j}^{\beta}}{\sum_{k=1}^{c} \tau_{i,k}^{\alpha} \times \eta_{i,k}^{\beta}} \tag{1}$$

Where β is the heuristic coefficient, α is the pheromone coefficient, $\tau_{i,j}$ is the pheromone value, $\eta_{i,j}$ is the maximizing contribution to the overall score, and c is the set of utilizable components.

For each solution, local pheromone is updated using the following equation.

$$\tau_{i,j} \leftarrow (1-\sigma) \times \tau_{i,j} + \sigma \times \tau_{i,j}^{0} \tag{2}$$

Where σ is the pheromone factor, $\tau_{i,j}^{0}$ is the initial pheromone, and $\tau_{i,j}$ is the pheromone for the graph edge (i, j).

Using the best candidate solution at the end of each iteration, pheromone is updated and decayed using the following equation.

$$\tau_{i,j} \leftarrow (1-\rho)\times\tau_{i,j} + \rho\times\Delta_{\tau_{i,j}} \tag{3}$$

Where $\Delta_{\tau_{i,j}}$ represents the maximizing cost for the best solution if the component i, j known otherwise it is 0, ρ represents the decay factor, and $\tau_{i,j}$ represents the pheromone for the graph edge (i, j).

3.3.2.Bees Colony Optimization

Bees Colony Optimization algorithm is inspired by the behaviour of bees in the real world. There are two types of bees namely scout bees and worker bees. Scout bees are sent to find out to search for the flower patches. On returning to hive, scouts communicate the information with the worker bees about the location and quality of the food source via a wangle dance. Then some of the scout bees return to the flower patch along with the worker bees and some scout bees go in search of new flower patches. And every time returning from the food source they communicate the quality of the food source.

The behaviour of bees is simulated to deal with the real world combinatorial problems. The basic idea of Bees Colony Optimization algorithm is to locate the good sites and explore them within the problem search space. Scouts are used to sample the problem search space and locate the good sites. These good sites are exploited using the local search where small number of good sites is explored compared to other sites. The pseudo code for the Bees Colony Optimization is given as follows.

```
Input:  Problem_size, Bees_no, Sites_no, Elite_Sites_no, Patch_Size_init, Elite_Bees_no,
        Other_Bees_no

Output: Bee_best

Population ← Initialize_Population(Bees_no, Problem_size)
while(Stop_Condition)
        Evaluate_Population(Population)
        Bee_best ← Get_Best_solution(Population)
        Next_Generation ← ( )
        Patch_size ← (Patch_size * Patch_Decrease_factor)
        Sites_best ← Select_Best_Sites(Population, Sites_no)
        foreach Sites_i ∈ Sites_best do
                Recruited_Bees_no ← ( )
                If (i < Elite_Sites_no)
                        Recruited_Bees_no ← Elite_Sites_no
                Else
                        Recruited_Bees_no ← Other_Bees_no
                End If
                Neighbourhood ← ( )
                For j to Recruited_Bees_no do
                        Neighbourhood ← Create_Neighbourhood_Bee(Sites_i, Patch_size)
                End For
                Next_Generation ← Get_Best_Solution(Neighbourhood)
        End foreach
```

Remaining_Bees_no ← (Bees_no – Sites_no)
For j to Ramaining_Bees_no do
Next_Generation ← Create_Random_Bee( )
End For
Population ← Next_Generation
End while
Return Bee_best

3.3.3.Simulated Annealing

Simulated Annealing is a global optimization technique which belongs to the field of Metaheuristic and Stochastic optimization. It is mainly used for function optimization and is an adaptation of Metropolis-Hastings Monte Carlo algorithm.

Simulated Annealing inspired by the annealing process of metallurgy. To increase the strength and durability of the material, first the material is heated so that atoms of the material can move freely and then the material is cooled slowly under the controlled environment. This process increases the size of the crystals in the material and condenses their defects. The main objective of this mechanism is to locate the minimum cost configuration in the problem search space. The pseudo code for the Simulated Annealing is given as follows

Input: Problem_Size, Max_no_Iterations, Max_Temperature

Output: Best_Solution

Current_Solution ←Create_Initial_Solution(Problem_Size)
Best_Solution ← Current_Solution
For i=1 to Max_no_Iterations do
Solution_i ←Create_Neighbour_Solution(Current_Solution)
Current_Temperature ←Calculate_Temperature(i, Max_Temperature)
If (Cost_Solution_i ≤ Cost_Current_Solution)
Current_Solution ←Solution_i
If (Cost_Solution_i ≤ Cost_Best_Solution)
Best_Solution ←Solution_i
End If
Else If $Exp\left(\frac{Cost\_Current\_Solution - Cost\_Solution\_i}{Current\_Temperature}\right) > Random(\ )$
Current_Solution ← Solution_i
End If
End For
Return Best_Solution

4. EXPERIMENTAL RESULTS

We have implemented the proposed model using C#. We have used publicly available Intel Berkeley Research lab dataset [8] for the experiment purposes. This dataset consists of three tables namely location table, aggregate connectivity strength table, and sensor data. Location table contains the information about the physical location of each sensor node. Aggregate connectivity strength table contains information about signal strength between each pair of sensor nodes. And sensor data table consists of 2.3 million readings with attributes date and time, epoch, sensor id, temperature, humidity, light, voltage.

The proposed model consists of three main modules. (i) Spatiotemporal relation model identification of correlation among sensors in both spatial and temporal dimensions by exploiting the DBSCAN algorithm. (ii) Selection of representative sensors. (iii) Finding the optimal transmission path data transmission using nature inspired algorithms namely Ant Colony Optimization, Bees Colony Optimization, and Simulated Annealing algorithm.

After applying DBSCAN, clusters were formed. Using the measurement strategy as explained in [18-19] we identified 23 sensor nodes as potential representative sensors. The figure 1 shows all the representative sensor nodes and the sensor nodes which they represent. The sensor nodes which are marked in a purple rectangle are the representative sensors and the boundary has been drawn if there is any sensor which is correlated to the representative sensor.

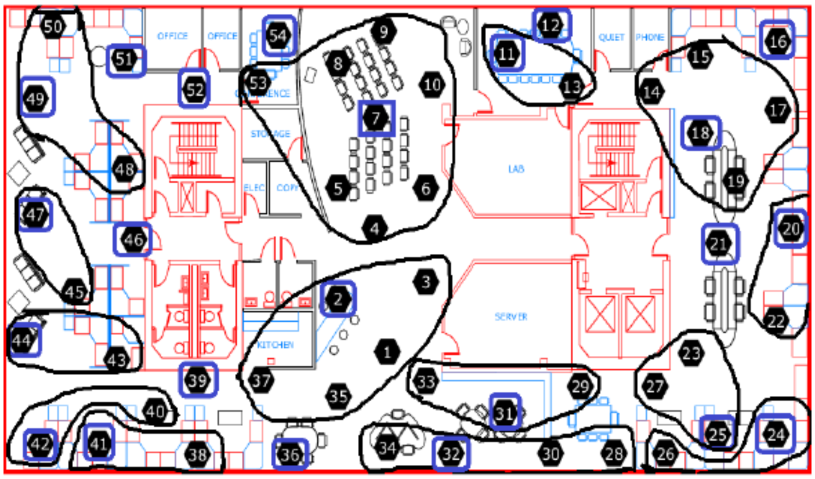

Figure 1 Representative Sensors

The optimal transmission path for data transmission is calculated using Ant Colony Optimization, Bees Colony Optimization, and Simulated Annealing algorithms. To calculate the best transmission path we have used the signal strength between each pair of sensor which is given in the aggregate connectivity strength table. The path is chosen as optimal path whose connectivity strength is higher compared to the other paths.

Table 1 gives the information about the parameters set in Ant Colony Optimization algorithm and figure 2 shows the output of the program. The optimal path found through the ACO is {47→2→12→42→46→44→52→31→20→25→24→49→39→41→36→54→11→18→21→16→51→7}

Table 1. Parameters of Ant Colony Optimization

Total number of ants used	10
Maximum execution time	1000
Pheromone influence α	4
Local node influence β	3
Pheromone evaporation coefficient (*rho*)	0.05
Pheromone increase factor (*Q*)	1

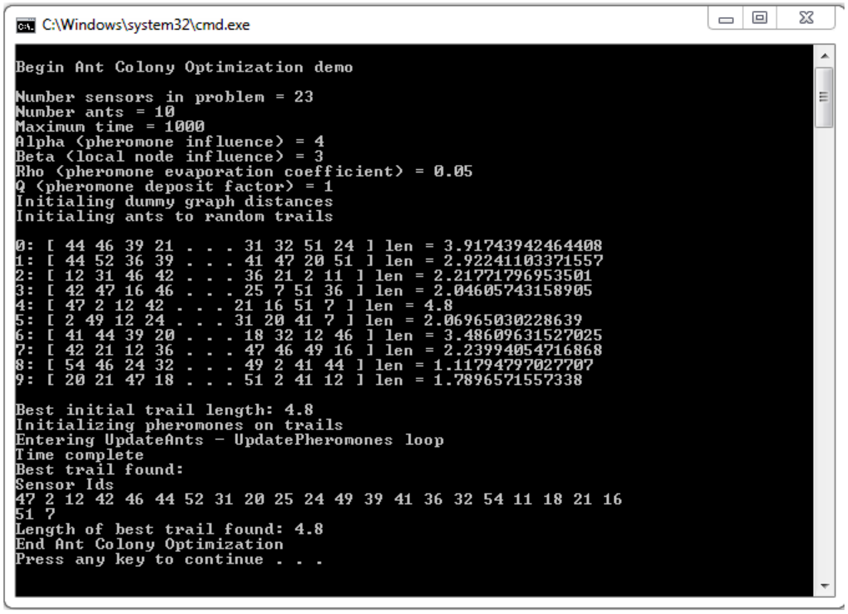

Figure 2 Output of ACO

Table 2 Parameters of Bees Colony Optimization

Total number of inactive bees	25
Total number of active bees	50
Total number of scout bees	25
Maximum number of visits	100
Maximum number of cycles	3460

Table 2 gives the details of the various parameters set in Bees Colony Optimization algorithm and the obtained result is shown in figure 3. The optimal path found through BCO is {44→47→7→41→21→42→52→12→54→36→39→2→11→46→24→18→16→32→20→51→31→49→25}.

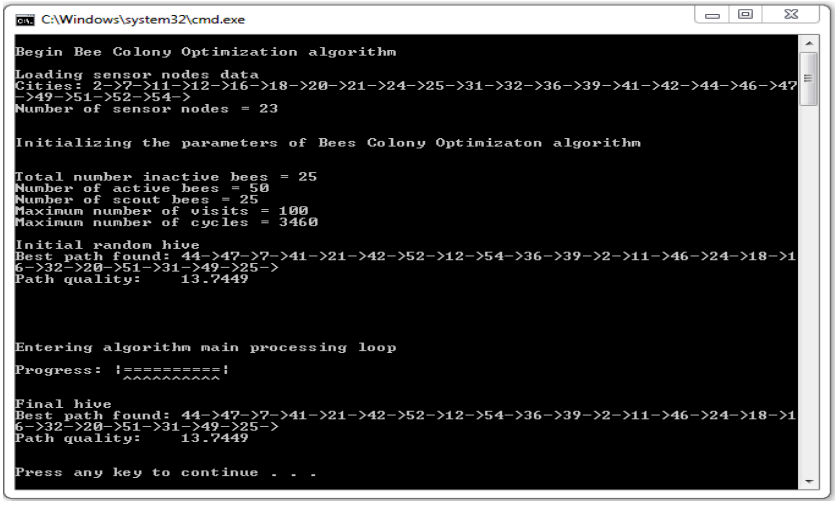

Figure 3 Output of Bees Colony Optimization

Table 3 Parameters of Simulated Annealing

Maximum Temperature	1000
Delta distance	0
Cooling rate	0.9999
Absolute Temperature	0.00001

Table 3 gives the details of various parameters values set in our experiment and figure 4 shows the result obtained. The optimal path found by the Simulated Annealing algorithm is {2→7→12→11→16→18→21→20→24→25→31→52→49→41→36→51→44→46→47→42→39→32→54}.

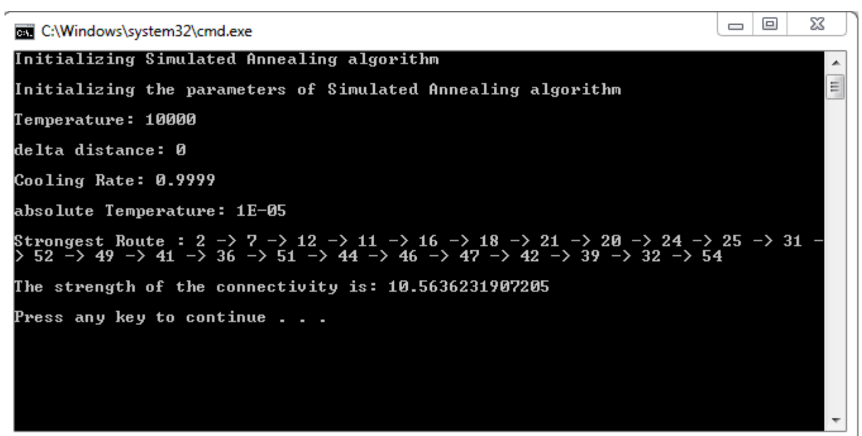

Figure 4 Output of Simulated Annealing

5. CONCLUSIONS

Developing an energy efficient model for sensor network has been a major research challenge and various methodologies have been proposed in the past. In our proposed model we have exploited the DBSCAN clustering technique to explore the correlation among the sensors in both spatial and temporal dimension. Once the clusters are formed we identified the subset of sensor nodes which best represents the entire network state. We have identified 23 representative sensor nodes which can be used to answer the queries instead of querying all 54 sensor nodes. This results in reduction in data acquiring and transmission of data by 57.40%. In the final stage we used nature inspired algorithms such as Ant Colony Optimization, Bees Colony Optimization, and Simulated Annealing techniques to find the optimal path for the data transmission. Finding an optimal path is necessary because if the path is not efficient or optimal then there is a possibility that data may not reach the destination successfully which result in retransmission of data causing more energy consumption. The optimal path increases the possibility of data being transmitted to the base station successfully. This results in decrease in the possibility of any packet being lost which results in reduction in consumption of energy.

The proposed model is best suited for static networks. In future we wish to address the key issue such as development of energy saving model for a dynamic sensor network.

REFERENCES

[1] J. Gehrke, S. Madden, (2004) "Query processing in sensor networks", *IEEE Pervasive Computing*. Vol. 3, No. 1, pp46-55.

[2] T. He. S. Krishnamurthy, J. A. Stankovic, T. Abdelzaher, L. Luo, R. Stoleru, T. Yan, L. Gu, J. Hui, B. Krogh, (2004) "Energy-efficient surveillance system using wireless sensor networks", in *Proceedings of the 2nd international conference on Mobile systems, applications, and services*. pp270-283.

[3] R. Szewczyk, E. Osterweil, J. Polastre, M. Hamilton, A. Mainwaring, D. Estrin, (2004) "Habitat monitoring with sensor networks", *Communications of the ACM*. Vol. 47, no. 6, pp34–40.

[4] S. Madden, M. J. Franklin, J. M. Hellerstein, W. Hong, (2003) "The design of an acquisitional query processor for sensor networks", in *Proceedings of the 2003 ACM SIGMOD international conference on Management of data*. ACM, pp 491–502.

[5] Y. Yao, J. Gehrke, (2003) "Query processing in sensor networks", in: *CIDR*. ACM, pp233-244.

[6] A. Deshpande, C. Guestrin, S. R. Madden, J. M. Hellerstein, W. Hong, (2004) "Model-driven data acquisition in sensor networks", in *Proceedings of the thirtieth international conference on Very large data bases.* VLDB Endowment, pp588–599.

[7] S. S. Skiena, (2008) "The algorithm design manual".

[8] Intel Berkeley Research lab dataset, http://db.csail.mit.edu/labdata/labdata.html , Accessed on 10.02.2012.

[9] A. Deshpande, C. Guestrin, S. Madden, (2005) "Using probabilistic models for data management in acquisitional environments", in *CIDR*. pp 317–328.

[10] D. Chu, A. Deshpande, J. M. Hellerstein, W. Hong, (2006) "Approximate data collection in sensor networks using probabilistic models", in *Data Engineering, 2006. ICDE'06 Proceedings of the 22nd International Conference on.* IEEE, pp48.

[11] D. Tulone, S. Madden, (2006) "Paq: time series forecasting for approximate query answering in sensor networks", in *Wireless Sensor Networks*. Springer, pp 21–37

[12] Y. Kotidis, (2005) "Snapshot queries: Towards data-centric sensor networks", in *Data Engineering, 2005. ICDE 2005. Proceedings. 21st International Conference on*. IEEE, pp131–142.

[13] A. Silberstein, R. Braynard, J. Yang, (2006) "Constraint chaining: on energy-efficient continuous monitoring in sensor networks", in *Proceedings of the 2006 ACM SIGMOD international conference on Management of data*. ACM, pp 157–168.

[14] S. Goel, T. Imielinski, (2001) "Prediction-based monitoring in sensor networks: taking lessons from mpeg", *ACM SIGCOMM Computer Communication Review*. Vol. 31, no. 5, pp82–98.

[15] W. R. Heinzelman, A. Chandrakasan, H. Balakrishnan, (2000) "Energy-efficient communication protocol for wireless microsensor networks", in *System Sciences, 2000, Proceedings of the 33rd Hawaii international conference on*. IEEE, Vol. 8.

[16] S. Yoon, C. Shahabi, (2007) "The clustered aggregation (cag) technique leveraging spatial and temporal correlations in wireless sensor networks", *ACM Transactions on Sensor Networks*. Vol. 3, no. 1, pp3.

[17] J. Han, M. Kamber, M, (2006) "Data Mining concepts and techniques". Morgan Kaufmann Publishers.

[18] E. Baralis, T. Cerquitelli, V. D'Elia, (2007) "Modeling a sensor network by means of clustering", in *Proceedings of the 18th international workshop on Database and Expert Systems Applications*. pp177-181.

[19] D. Apiletti, E. Baralis, T. Cerquitelli, (2011) "Energy saving models for wireless sensor networks", *Knowledge and Information Systems*. Vol. 28, pp615-644.

[20] J. Brownlee, (2011) "Clever Algorithms Nature Inspired Programming Recipes".

[21] M. Dorigo, V. Maniezzo, A. Colorni, (1991) "Positive feedback as a search strategy", Technical. Report.

[22] M. Dorigo, V. Maniezzo, A. Colorni, (1996) "Ant system optimization by a colony of cooperating agents", *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*. IEEE, Vol. 26, no. 1, pp29–41.

Authors

Doreswamy received B.Sc and M.Sc Degree in Computer Science from University of Mysore in 1993 and 1995 respectively. Received Ph.D degree in Computer Science from Mangalore University in the year 2007. After completion of his Post-Graduation Degree, he subsequently joined and served as Lecturer in Computer Science at St. Joseph's College, Bangalore from 1996-1999.Then he has elevated to the position Reader in Computer Science at Mangalore University in year 2003. He was the Chairman of the Department of Post-Graduate Studies and research in computer science from 2003-2005 and from 2008-2009 and served at varies capacities in Mangalore University. At present he is the Chairman of Board of Studies and Professor in Computer Science of Mangalore University. His areas of Research interests include Data Mining and Knowledge Discovery, Artificial Intelligence and Expert Systems, Bioinformatics, Molecular modelling and simulation, Computational Intelligence, Nanotechnology, Image Processing and Pattern recognition. He has been granted a Major Research project entitled "Scientific Knowledge Discovery Systems (SKDS) for Advanced Engineering Materials Design Applications" from the funding agency University Grant Commission, New Delhi, India. He has published about 30 peer reviewed Papers at national/International Journal and Conferences. He received SHIKSHA RATTAN PURASKAR for his outstanding achievements in the year 2009 and RASTRIYA VIDYA SARASWATHI AWARD for outstanding achievement in chosen field of activity in the year 2010.

Srinivas Narasegouda received his BSc degree from Gulbarga University in 2006 and MSc degree from Karnatak University in 2008. At present, he is pursuing his PhD under the guidance of Professor Doreswamy in the department of computer science Mangalore University, Mangalore. His areas of interest are Data Mining and Knowledge Discovery, Swarm Intelligence, and Symbolic Data Analysis.