ISSN: 2321-8169
2208 – 2212

Data Density Correlation Degree Clustering Algorithm for Multiple Correlated Sensor Networks using Fuzzy Logic

Priyanka Sonar
Dept. of Computer Engineering
D. Y. Patil College of Engineering-Akurdi,
Pune, India.
priya.r.sonar@gmail.com

Ms.Shanti Kumarguru
Dept. of Computer Engineering
D. Y. Patil College of Engineering-Akurdi,
Pune, India.
gurushaanguru@gmail.com

Abstract— Wireless Sensor Network (WSN) is the collection of several sensor nodes which are able to send their sensed data to base station. In dense WSN, consecutive observations obtained by sensors are spatial as well as temporally correlated in applications that involve the observation of the variation of a physical phenomenon. These sensor nodes are battery driven, therefore an efficient utilization of power is essential in order to reduce data traffic inside sensor networks and thus reduce amount of data that need to send to base station for enhancing the network lifetime. For this reason data aggregation is used. The correlation degree gives a correlation measurement that measures the correlation between a sensor node's data and its neighboring sensor nodes' data. The resulting representative data obtained using the proposed methods have a lower data distortion than those obtained earlier. Also, to construct an energy balanced network in data transmitting process, the energy of every sensor nodes should be considered and Fuzzy Logic is also used to determine an optimal routing path from the source to the destination by maintaining the highest remaining battery power, minimum traffic loads and minimum number of hops.

Keywords- *Wireless Sensor Networ, Spatial Correlation, Temporal Correlation, Data Density Correlation Degree, Fuzzy Approach*

*****

I. INTRODUCTION

Wireless Sensor Networks (WSNs) are the collection of number of small nodes for performing some functions as processing, gathering sensory information and to communicate with other connected nodes in the network. Features of WSN are homogeneous devices, no wired infrastructure, potential multi hop routes, dispersed large network size and self-organization. The type nodes may be mobile or stationary and all nodes act as routers. Intermediate nodes are responsible to establish communication between two unconnected nodes [2]. Every node in network that falls inside the communication range of a node is considered to be reachable from every node. WSN have many applications like Industrial control and monitoring, Home automation, Security, Military sensing and directing, Consumer electronics, Asset tracking and supply chain management, Intelligent Agriculture, commercial applications, Fire alarming, Environmental monitoring and health monitoring etc.

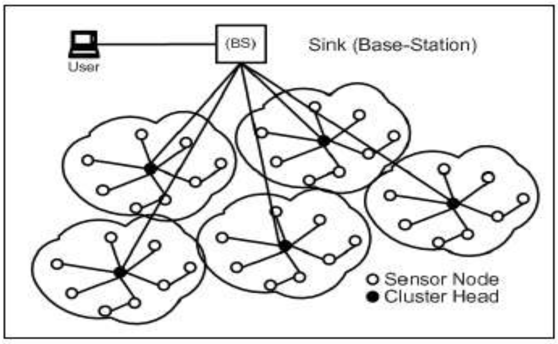


Figure 1. Sensor network architecture [1]

The architecture of a cluster based WSN is shown in Fig 1. All the sensor nodes send their sensed data towards the CH which aggregate and then forward only the meaningful data to the sink or base station (BS). The function of sink is to collect the data from all the CH which transmits the data to the user via Internet or satellite. Each sensor node makes its own decisions according to its mission, knowledge of its computing, the information it currently has and available energy resources.

This paper proposes the Data Density Correlation Degree clustering algorithm for multiple correlated sensor nodes, which efficiently resolve the problems in recent correlation model. Also, the generated cluster shape can be easily adapted to the environment.

II. RELATED WORK

Many researches were given to study the correlation in WSN in recent years. It gives the information about the theoretical aspects of the correlation. To increase lifetime of WSN is an important aspect of intensive research [8]. Many methods are suggested in order to minimize the nodes energy consumption. Some different energy-efficient techniques in data aggregation are as follows:

A. *Grid-Based Architecture*

This scheme is used for energy efficient data storage [10]. It gives a Snake-like Energy Efficient Scheduling and network is divided into 2 dimensional logical grids where the number of sensors in a grid is N. It basically works on Active and sleep mode procedure i.e. if one sensor is active at one time slot then the other is in sleep mode at that time of slot.

B. *Steiner Tree*

This technique [11] describe the WSN communication model as an undirected graph G < V; E >, where V is a node set contains all nodes that are present within the region, E is the distance between nodes in the graph associate with those vertices. Communication distance of the node is R. It gives the introduction of a Steiner tree, a weighted undirected graph G is given, and the demand Steiner tree T and find the shortest path from the root of T to other nodes.

C. *Polynomial Regression Based Secure Data Aggregation*

In this technique, the author suggests a protocol in which sensor nodes represents their sensed data as polynomial functions [12]. This paper proposes a Polynomial Regression based secure Data aggregation protocol, called PRDA, used to preserve the privacy of the data being aggregated. PRDA is nothing but a additive data aggregation protocol and basically achieves data privacy by employing polynomial regression on sensor data series. The main idea behind PRDA protocol is to perform data aggregation using polynomial coefficients and these are used to represent sensor data.

D. *Clustering Based Lifetime Maximizing Aggregation Tree*

In this technique [13], author created aggregation tree which is then used to reduce energy consumption. It minimizes the distance traversed. In this technique the node having maximum available energy is considered as parent node or aggregator node. It creates best possible aggregation tree minimizing energy utilization, minimizing cost and hence as a result maximizing network lifetime. Therefore by achieving these parameters the obtained aggregation tree proves to be the best for enhancing the network lifetime.

III. DATA DENSITY MULTIPLE CORRELATION DEGREE CLUSTERING ALGORITHM

The proposed correlation degree is a multiple correlation measurement including spatial as well as temporal that measures the correlation between a sensor node's data and its neighboring sensor nodes' data over the time. In typical WSN applications, to achieve satisfactory coverage, sensor nodes were densely deployed [3]. Therefore, when a single event occurs in the sensor field at that time multiple sensor records same information. If there is high density in the network topology, these sensor observations are highly correlated. The degree of correlation increases with decreasing internodes separation. In some of the WSN applications such as event tracking there is a need for sensor nodes to periodically perform observation and then transmission of that sensed event features. This nature of the energy-radiating physical phenomenon creates the temporal correlation between each consecutive observation of a sensor node [4]. The degree of correlation between consecutive sensor measurements may vary according to the temporal variation characteristics of the phenomenon.

IV. FUZZY APPROACH

Fuzzy model is the energy-aware routing in WSN. Many routing protocols uses fixed or crisp metrics to make energy-aware routing decisions, but these protocols are not easily adapted to changes in sensor types because according to the type of sensor node implementation platform energy metrics vary widely. Also, some of the factors for calculating routing metric are conflicting. Therefore, fuzzy logic has a great potential to deal with conflicting situations and have a great imprecision in data using heuristic human reasoning without needing complex mathematical modeling. Fuzzy approach is used to give a tunable parameter based approach by using fuzzy variables and rule base. As a result, WSNs are capable of handling a wide range of energy metrics of different sensor implementation platform.

In fuzzy systems, the dynamic behavior of a system is characterized by a set of fuzzy rules mainly based on the knowledge of a human expert. These rules are of the general form IF antecedent(s) THEN *consequent*(s), where antecedents and consequents both are propositions containing different variables in set. Antecedents of a fuzzy rule form a combination of fuzzy sets by using logic operations [12]. Thus, fuzzy sets and fuzzy rules combines together to form the knowledge base of a rule-based inference system. Rules are very important aspect of a fuzzy system. These rules are provided by experts or can be extracted from numerical data. Also, the rules that we are interested in can be expressed as a collection of IF THEN statements in fuzzy approach.

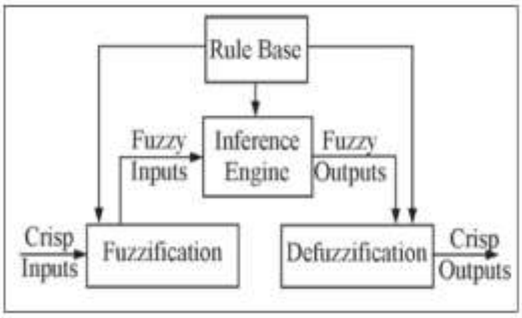


Figure 2. Typical structure of the fuzzy approach.

Fig. 2 shows the typical structure of a fuzzy system. It mainly consists of four components as; fuzzification, rule base, inference engine and defuzzification.

V. IMPLIMENTATION DETAILS

A. *Design*

The WSN is modeled by undirected graph G = (V , E). Where V is the set of all sensor nodes in the WSN, E is the edge set consisting of all links in the WSN. The antenna of sensor node i ($i \in$ V) is an omnidirectional antenna, with a communication radius of $\alpha(i)$. Let N(i) be the set of sensor nodes within the circle of the communication radius of i . In cluster-based data aggregation networks, every cluster head sends aggregated data obtained from its member nodes to the sink node by one hop or multi-hops. Fig.3 shows the block architecture of proposed system.

B. *Implimentation of Join Spatial-Temporal correlation on extended DDCD Clustering Method*

In a WSN, if a certain number of neighboring sensor nodes' data are close to a sensor node's data at a specific time, this sensor node can represent its neighbors in the data domain. This representative sensor node is called the core sensor node(CSR). The definition of CSR can be given as, "Assume sensor node v has n neighboring sensor nodes. They are respectively v_1, v_2,…., v_n. The data object of v is D. Its neighboring sensor nodes' data objects are respectively

ISSN: 2321-8169
2208 – 2212

$D_1, D_2, \ldots, D_n$. If there are N data objects in $D_1, D_2, \ldots, D_n$ whose distances to D are less than ε and $minPts \le N \le n$ then the sensor node v is called the core sensor node. Where *minPts* is the amount threshold, ε is the data threshold. "

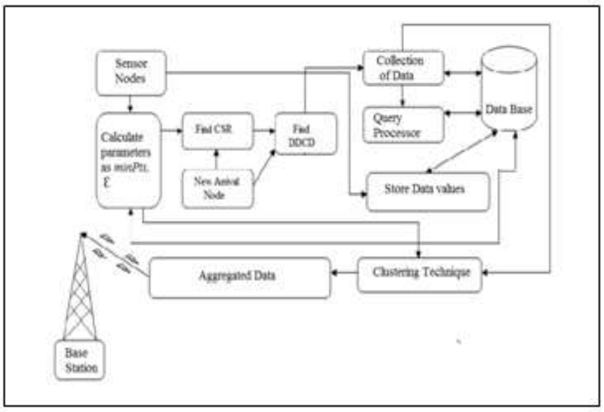


Figure 3. Architecture of Proposed System

Suppose sensor node v has total n number of neighboring sensor nodes which are within the cycle of the communication radius of v. They are $v_1, v_2, \ldots, v_n$, respectively. The data object of v is D, and its neighboring sensor nodes' data respectively $D_1, D_2, \ldots, D_n$. Among these n data objects, there are N data objects whose distances to D are less than ε, and $minPts \le N \le n$. Then the data density correlation degree of sensor node v to the sensor nodes whose data objects are in ε-neighborhood of D is as follows.

$$Sim(v) = \begin{cases} 0, & N < minPts \\ a_1\left(1 - \frac{1}{\exp(N - minPts)}\right) + a_2\left(1 - \frac{d_\Delta}{\varepsilon}\right) + a_3\left(1 - \frac{d}{\varepsilon}\right), & N \le minPts \end{cases} \quad (1)$$

Where *minPts* is the amount threshold, ε is the data threshold, d∆ is the distance between D and the data center of the data objects which are in the ε-neighborhood of D. d is the average distance between the N data objects and D. a_1, a_2 and a_3 are weights. $a_1 + a_2 + a_3 = 1$.

The proposed clustering algorithm includes three procedures: the Sensor Type Calculation (STC) procedure, the Local Cluster Construction (LCC) procedure and Global Representative sensor node Selection (GRS) procedure. In the STC procedure, each sensor node judges itself whether it is a core sensor node according to definition, and its data density correlation degree is calculated by Eq.1. Meanwhile, it records the sensor nodes that are in the ε-neighborhood of its data and those which are not in the ε-neighborhood. In the LCC procedure, local clusters are constructed using the results achieved in the STC procedure. In the GRS procedure, local clusters are merged according to the maximal DDCD information stored in every sensor node, and a representative sensor node is selected in each merged cluster. As these three procedures are accomplished in a WSN, sensor nodes are classified into three types: representative sensor nodes (RSN), isolated sensor nodes (ISN) and member sensor nodes (MSN). The RSN and ISN are responsible for sensing and sending sampled data to the sink node. While the MSN just transmit sampled data collected by RSN or ISN, or do nothing.

C. Implimentation of Fuzzy Approach

To use a fuzzy approach, the fuzzified values are processed by the inference engine, which consists of a rule base and various methods to inference the rules [13]. The rule base is simply a series of IF-THEN rules that relate the input fuzzy variables and the output variable using linguistic variables each of which is described by fuzzy set and fuzzy implication operator AND. Table I shows the IF-THEN rules used in the proposed method. As an example, IF *RE(n)* is *very high* and *TL(n)* is *very* low THEN NC(n) is very high. All these rules are processed in a parallel manner by a fuzzy inference engine. Any rule can definitely fires contribute to the final fuzzy solution space. At the end, the defuzzification finds a single crisp output value from the solution fuzzy space. This value represents the node cost [11].

No	Antecedent		Consequent
	RE(n)	*TL(n)*	*NC(n)*
1	VL	VL	L
2	VL	L	VL
3	VL	M	VL
4	VL	H	VL
5	VL	VH	VL
6	L	VL	M
7	L	L	M
8	L	M	L
9	L	H	L
10	L	VH	VL
11	M	VL	H
12	M	L	M
13	M	M	M
14	M	H	L
15	M	VH	L
16	H	VL	VH
17	H	L	H
18	H	M	H
19	H	H	M
20	H	VH	M
21	VH	VL	VH
22	VH	L	VH
23	VH	M	VH
24	VH	H	H
25	VH	VH	H

D. Mathematical Model

A mathematical model of the proposed system is given below.

1. Total system S is defined as
   S = {
2. Identify input as
   I = { ε ; *minPts*; a_1; a_2; a_3; *N(i)* }
   ε = Data Threshold
   minPts = Amount Threshold
   a1; a2; a3 = Weights
   N(i) = Neighboring Sensor nodes set of Sensor Node *i*
3. Identify output as
   O = {R}
   R = Representative Sensor Nodes

 ISSN: 2321-8169
2208 – 2212

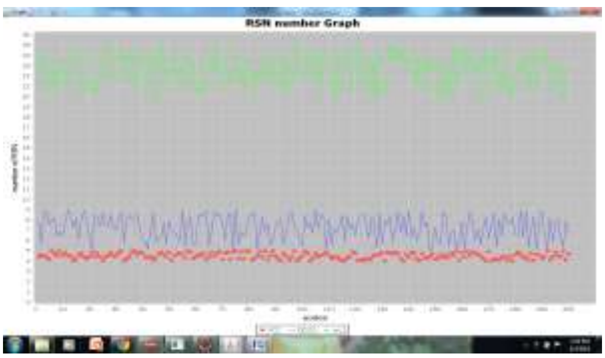
4. Identify the sensor nodes v has n neighboring sensor nodes
   $V = v_1, v_2, \ldots, v_n$
5. Identify core sensor nodes as *CS* and Non-core sensor nodes as *NCS*.
6. Identify the data object of v is D. Its neighboring sensor nodes' data objects are respectively
   $D = D_1, D_2, \ldots, D_n$.
7. By using (1) identify the data density correlation degree of sensor node v to the sensor nodes whose data objects are in ε - neighborhood of D.
8. $NodeSet_{inner}(i)$ includes the IDs of the sensor nodes whose data are in the ε - neighborhood of the data of sensor node i.
9. $NodeSet_{outer}(i)$ includes the IDs of the sensor nodes whose data are not in the ε - neighborhood of the data of sensor node *i*.
10. Identify *Sim(i)* as Data Density Correlation Degree.
11. Apply the rule base series of IF-THEN rules that relate the input fuzzy variables and the output variable using linguistic variables each of which is described by fuzzy set and fuzzy implication operator AND.

VI. RESULTS

A. Data Set

A dataset is a collection of data. The result of DDCD Clustering method gives clustering between nodes and energy consumption chart. A simple simulation is performed on 200 sensor nodes. Out of these 200 sensor nodes, 5 are Source sensor, 3 are Receiver sensors and 2 are Attacker nodes. The simulation is executed for 00:05:50 minutes. After that results occurred are as follows.

(a)



(b)

Figure 4. Clustering performance comparison results. 4(a), 4(b) and 4(c) Global average relative errors, numbers of RSN and numbers of ISN for different clustering methods, respectively.

(c)

This system is related to some of the WSN applications such as event tracking may require sensor nodes to periodically perform observation and transmission of the sensed event features.

VII. CONCLUSION

Now a days, to maximize the lifetime of WSN, especially when they have a limited energy supply, is becoming a great challenge. Power management and energy-efficient communication techniques are very necessary, to extend the network lifetime. So the proposed method introduces extended DDCD clustering method along with fuzzy approach. This method is proposed to measure the spatial correlation as well as temporal correlation between sampled data. In the DDCD clustering method, sensor nodes in the same cluster have a high correlation degree, while those belonging to different clusters have a low correlation degree. The performance of this clustering method is more energy efficient and could obtain better data representation than the other clustering methods. Thus, DDCD clustering method is useful for the application where the sensor nodes are densely deployed and the sampled data change slowly with time.

The direct towards further work, in data transmitting process, mobile sensor nodes will be considered to construct the more energy balanced networks.

ACKNOWLEDGMENT

The authors would like to thank the publishers, researchers for making their resources available and teachers for their guidance. We also thank the college authority for providing the required infrastructure and support. Finally, we would like to extend a heartfelt gratitude to friends and family members.

REFERENCES

[1] Manjunatha, P.; Verma, A.K.; Srividya, A., (2008) "Multi-Sensor Data Fusion in Cluster based Wireless Sensor Networks Using Fuzzy Logic Method," Industrial and Information Systems., IEEE Region 10 and the Third international Conference on, vol., no., pp.1,6, 8-10 Dec.

[2] N. Bulusu, D. Estrin, and L. Girod, "Scalable Coordination for Wireless Sensor networks: self-configuring localization

ISSN: 2321-8169
2208 – 2212

systems," Proc. of the 6th Symposium on Communication Theory and Applications(ISCTA'01), Ambleside, UK,July 2001

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," Computer Networks (Elsevier) Journal, vol. 38, no. 4, pp.393-422, March 2002.

[4] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: A survey," IEEE Wireless Commun., vol. 11, no. 6, pp. 6–28, Dec. 2004.

[5] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor network," IEEE Trans. Wireless Commun., vol. 1, pp. 660–670, 2002..

[6] T. A. Runkler, "Selection of appropriate defuzzification methods using application specific properties," IEEE Trans. Fuzzy Syst., vol. 5, no. 1, pp. 72–79, Feb. 1997.

[7] Imad S. AlShawi, Lianshan Yan, Senior Member, Wei Pan, and Bin Luo, Member, IEEE, "Lifetime Enhancement in Wireless Sensor Networks Using Fuzzy Approach and A-Star Algorithm," Comput. Netw., vol. 45, no. 3, pp. 245–259, 2004.

[8] Jingsha He, Mingxin Yang , Xuguang Sun, Research On Data Aggregation in Wireless Sensor Networks Based on Clustering Method, International Technology and Applications, IEEE Conference, 2011.

[9] K.-Y. Cai and L. Zhang, "Fuzzy reasoning as a control problem," IEEE Trans. Fuzzy Syst., vol. 16, no. 3, pp. 600–614, Jun. 2008.

[10] R. V. Kulkarni, A. Forster, and G. K. Venayagamoorth, Computational Intelligence in Wireless Sensor Networks: A survey IEEE Commun.

[11] Surveys Tutorials, vol. 13, no. 1, pp. 68-96, Feb. 2011.

[12] Imad S. AlShawi, Lianshan Yan, Senior Member, Wei Pan, and Bin Luo, Member, IEEE, "Lifetime Enhancement in Wireless Sensor Networks Using Fuzzy Approach and A-Star Algorithm," Comput. Netw., vol. 45, no. 3, pp. 245–259, 2004.

[13] K.-Y. Cai and L. Zhang, "Fuzzy reasoning as a control problem," IEEE Trans. Fuzzy Syst., vol. 16, no. 3, pp. 600–614, Jun. 2008.