Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 5, Issue. 1, January 2016, pg.121 – 126

SPECTRAL CLASSIFICATION (SCRC) ALGORITHM FOR ENERGY SAVING IN WSN

Ms. Chaitra[1], Mrs. Nitty Sarah Alex[2]

[1]M.Tech Student, Department of Information Science and Engineering, New Horizon College of Engineering, Bangalore, India
[2]Senior Assistant Professor, Department of Information Science and Engineering, New Horizon College of Engineering, Bangalore, India
[1] chaitrajmr@gmail.com; [2] nittyalex82@yahoo.co

Abstract*— A wireless sensor network is comprised of self-organized sensor nodes which are distributed in the environment to collect, process and transmit the collected data from the environment. These sensor nodes are small in size and the mainly used in deploying applications based on WSNs that make use of the collected data from sensor nodes to raise the real-time decisions. Since sensor nodes are battery powered devices with limited processing and transmission power the main issue is energy consumption of WSNs. Energy saving to extend the wireless sensor network lifetime involves great challenges. So, we reviewed algorithm called as Spectral Classification for Robust Clustering in WSNs (SCRC-WSN) to understand the impact of node density with its energy and its lifetime gains. It was observed the lifetime of whole network is improved and presents more energy efficiency compared to the Low-Energy Adaptive Clustering Hierarchy (LEACH) approach and centralized LEACH.

***Keywords*— Spectral Classification, Clustering, Energy Consumption, Wireless sensor network**

I. Introduction

Wireless sensor network consisting of spatially distributed devices using sensors, these sensors used to monitor physical and environmental conditions. These sensor nodes are small in size and capable of sensing physical phenomenon and processing them. The applications are deploying based on WSNs is to make use of data sensed by sensors.

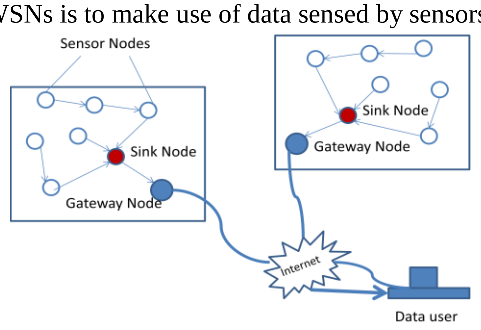

Figure1: Wireless Sensor Network

The fig 1. shows that the sensor nodes are distributed in the environment. Each sensor nodes can sense, process, talk to its peer so as to work together in cooperative manner. The sink node is wired to Gateway node which provides out-of-network connectivity. Gateway nodes manage the wireless network and it aggregates the data from distributed sensor nodes.

The various methods of WSNs used for extract information from a data set and transform it into an understandable structure for further use are data collection method, data fusion method, tracking method, etc. The main aim of data collection method [1] is to collect large amount of data and to reduce the data loss due to less memory capacity of sensor nodes. Data Fusion method [2] is generally defined as the use of techniques that combine data from multiple sources and gather this information in order to achieve inferences ,this will be more efficient and potentially more accurate than if they were achieved by means of a single source. In target tracking method [3] the sensor nodes which can sense target at particular time are kept in active mode while remaining nodes are to be retained in inactive mode so as to conserve energy until the target approaches them.

The main techniques used in WSN are Key management technique, Routing technique, Localization techniques, etc. WSN needs a cryptography algorithm that should be select carefully and the most important factor for those algorithms solving the key management agreement problems [4] that is needed to provide an encrypted and authenticated data transmission between the sensor nodes to have a secure channel. There are two routing techniques [4] single path routing technique and multipath routing technique. Multipath routing technique [5] is more efficient technique to route data in WSN because it can provide reliability, security and load balancing. The important function of sensor network is to collect data and send to the destination. It is important to know about the location of the data collected which can be done using localization technique [6]. The main goal of data aggregation technique [7] is to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced.

The major issues of WSN are: Data generated by WSNs is highly resource constrained, small in size, huge in volume , fast changing it is very challenging to design data mining techniques for WSNs. It is not feasible to store the entire WSNs data to a centralized location as it demands huge amount of memory and high storage cost. Centralized approach in WSN is difficult and often not suitable because of various reasons such as limited communication bandwidth and limited power supply for running sensor nodes.

Data mining is a part of knowledge discovery process which involves a process of extraction of useful information and patterns from huge data which is also called knowledge discovery process. Knowledge discovery process in databases comprises of a few steps leading from raw data collections to some form of new knowledge. This process consisting of following steps:

Data cleaning-It is a phase in which noise data and irrelevant data can be removed from the collection.

Data integration-at this phase multiple data sources, often heterogeneous combined in a common source. Data cleaning and data integrations can be performed together as a preprocessing phase to generate data warehouse.

Data selection-the data relevant to the analysis is decided on and retrieved from the data collection.

Data transformation-at this phase in which the selected data is transferred into forms appropriate for the mining procedure. **Data mining**-in this phase the clever techniques are applied to extract patterns potentially useful.

Pattern evaluation -strictly interesting patterns representing knowledge are identified based on given measures.

Knowledge representation- This is the final phase this step is used to visualization techniques to help to users understand and interpret the data mining results.

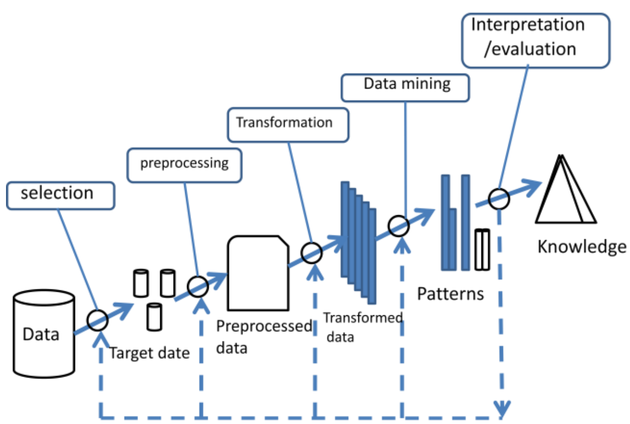

Figure2: Knowledge discovery process.

II. RELATED APPROACH

The main four types of data mining techniques used for WSNs are frequent pattern mining, sequential pattern mining, clustering, classification. All these techniques can be used in both centralized and distributed environment in WSNs.

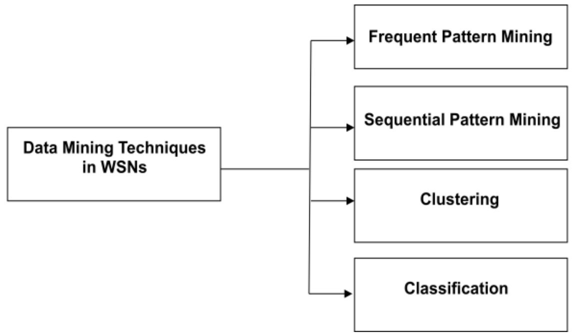

Figure3: Data mining techniques in WSNs

Frequent pattern mining: Finding patterns (item set, sequence, structure) that occur frequently in dataset. sequential that are relevant between data where the values are delivered in sequence**. Clustering** : process of partitioning set of data(or objects) into a set of meaningful subclasses called clusters. **Classification**: it is used to predict group membership for data instances.

Traditional data mining techniques used for sensor data transmission in WSNs have challenges in term of resource constraint, fast and huge data arrival, data transmission, dynamic network topology etc.

Existing traditional data mining techniques are mainly uses centralized approach. In this technique data is directly collected at the central location which is not restricted by computational resources. Also compared with the traditional datasets, in WSNs data in systems is continuous, varying and having different data rates .so, it is not feasible to store the entire WSNs data to centralized location as it demands huge amount of memory and high storage cost.

Using data mining techniques we can discover knowledge from collected data in wireless sensor networks automatically, intelligently to extract the information required by user. Sensor nodes in WSN are highly based on energy of sensor nodes. So, minimizing the energy consumption in sensor nodes is major issue in WSN. Sensor nodes are limited to battery power and memory space which requires a suitable data mining technique to minimize energy consumption.

III. PROBLEM DEFINITION

In wireless sensor networks sensor nodes are usually battery powered, but it is not practical to recharge or replace the batteries of all the sensors. Indeed, the number of sensor nodes in a given WSN, is too large on the other hand, these nodes are positioned in remote, battlefield, desert or hostile areas. Since sensor nodes are power constrained devices, frequent and long distance transmissions should be kept to minimum in order to extend the network lifetime. Moreover, in WSN, a large part of energy is consumed when the wireless communications are established. Therefore, direct communications between nodes and the base station are not encouraged. Due to these assessments and in order to enhance the network lifetime, particular innovative techniques that improve energy efficiency are highly required**.**

There are many techniques has been proposed such as LEACH and LEACH-C(centralized LEACH). LEACH algorithm [8] is a fundamental method in homogeneous clustering techniques and it chooses cluster heads periodically and distributes consumed energy uniformly by role rotation. Under the heterogeneous network, this protocol will become poor and not efficient.

In Centralized LEACH algorithm [9], the different cluster heads are elected by Base Station (BS). The later starts receiving all information about each node regarding its location and energy level. This algorithm is not feasible for large networks because nodes that are far away from the BS, will have problem sending their states to BS and as the role of cluster heads rotates all the time, the far nodes will not reach the BS in quick time increasing the latency and delay.

In order to overcome above problems the spectral methods for clustering have recently started to get a great attention in many research areas. These methods make use of the spectrum of the adjacency matrix of the data to cluster a considered set of elements. They are considered as powerful techniques in data analysis. For this reason, we propose to include the concepts to manage a given WSN and to improve its lifetime.

IV. PROPOSED ALGORITHM

In this paper, a new homogeneous clustering protocol for WSN is presented. It is based on the spectral clustering that included a variety of methods based on the notion of similarity matrix and using the eigenvectors differently. In the spectral method the data dividing into K disjoint classes based on the K eigenvectors related to K largest Eigen values of a Laplacian matrix. In this paper we consider a network with N nodes, uniformly distributed in a M×M square region. Moreover we assume that the network topology remains unchanged over time.

A. Clustering step

Each node has the global knowledge of the network which is collected in the base station with its node positions and applies the clustering process. Assume that each node knows its location which can be obtained at a low cost by a global positioning system or by using other localization systems. Then the WSN nodes transit their location in a short message to the BS.

Each WSN can be represented by its corresponding undirected graph G(V,E). where V is the set of vertices(nodes) representing different sensor nodes and E is the edge set enclosing all dependencies between the nodes. Each vertex of V is identified by an index i ϵ [1……N]. We assume that G is a graph without loops and multi edges.

Let A$\epsilon R^{N\times N}$ be the adjacency matrix of the graph G. Each value of A is associated to each pair of the graph nodes(i,j). This value is of Gaussian type and the matrix A is given by the equation

$$A = [a_{ij}] = \begin{cases} e^{\left(-\frac{1}{2\sigma^2} d^{2(i,j)}\right)} \\ 0 \end{cases}$$

The total weight of edges incident to node i is given by $d_{ii}=\sum_{j=1}^{N} a_{ij}$. The degree matrix D$\epsilon R^{N\times N}$ of G is diagonal matrix defined by D=[d_{ij}] and the N×N Laplacian matrix of the graph is defined by.

$$\mathbf{L} = \boldsymbol{D^{1/2}}\ \mathbf{A}\ \boldsymbol{D^{1/2}}$$

The objectives of the current step are to define the optimal number of clusters and to form them. Based on the Laplacian matrix L defined above, we form a new matrix U composed of the K eigenvectors related to K largest Eigen values of L. In order to determine the K clusters of WSN, we apply the classification algorithm k-means to the matrix U.

By considering K cluster, the energy depends on the distance between the cluster heads and the non cluster heads of each cluster i.e E_{Round} is minimal if the quantity:

$$\sum_{i=1}^{1} \mathrm{d}_{\mathrm{i}}^{4} \cdot E_{mp} + \sum_{j=1}^{N-K-1} d_j^2 \cdot E_{fn}$$

is minimal. The objective function that allows to decide whether to reconsider the partitioning process or not of the WSN, is defined by the distance matrix

$$M_{dis}{}^{k}\ (\ M_{dis}{}^{k}=[dis_{ij}{}^{k}];\text{with } dis_{ij}{}^{k}$$

is the distance between the node I and the node j of the cluster labeled k) of each cluster . The allowed threshold to this function is d_0. Hence if at least one element of any $M_{dis}{}^{k}$ is greater than d_0., the considered number of clusters will be reused. Otherwise, the optimal number of clusters is K.

B. Cluster head election step

Once the clusters are determined, the next step consists in defining the cluster heads where the numbered node id (identification) will be in some random position on the cluster. Thus the cluster head in each round of communication will be at a random position on the cluster. It is so important that nodes die at random locations of the network. The idea behind this is to make the sensor network robust to failures. Moreover, by taking in consideration the nodes id in clusters, the possible cluster heads will be determined. Indeed, in the round r of the simulation, we use the number c_k =(r mod $|S_k|$) to select the suitable cluster head for the appropriate cluster; where $|S_k|$ represents the total number of nodes in a defined cluster K. Besides, if the residual energy Eng_{ck} of the node, with id=ck, is greater than a threshold Θ_{Eng}, this node will be the cluster head of the cluster k in the round r. We define Θ_{Eng} as the minimum residual energy required for a given node to be cluster head. It is the summation of the energy needed to receive and process data coming from the appropriate cluster nodes, and to transmit towards the BS. This Θ_{Eng} is given below:

$$\Theta_{Eng} = L \cdot ((|S_k| + 1) \cdot E_{elec} + |S_k| . E_{DA} + E_{mp} . d_1^4)$$

$i \neq j$

Otherwise

Where d_i is the distance between the node i and the BS. If the residual energy Eng_{ck} of a possible cluster head is less than the threshold Θ_{Eng}, this node must send a short message informing the node with id=c_{k+1} to be the new possible cluster head at the iteration r, and so on.

Consequently, each cluster head will be able to collect data from the cluster nodes and will transmit the aggregate information to the BS. Thus, the number of the direct transmission is efficiently reduced and the whole network lifetime is extended. In addition, energy consumption will be distributed with more equitability between all nodes.

C. Data transmission

Once clusters and cluster heads are created, each cluster head knows which nodes it is supervising. Based on the node's id in the appropriate cluster, a Time Division Multiple Access (TDMA) MAC protocol schedule assignment will be generated automatically. If we suppose that the node with id = I is elected as a cluster head, the node with id = (i+1+$|S_k|$) mod $|S_k|$ will take the first time slot to transmit, where $|S_k|$ is the total number of nodes in the defined cluster K. Here, we avoid the techniques applied by the traditional algorithm which consume more energy and ask for more synchronization when the cluster heads are elected to assign the TDMA access. Moreover, this technique guarantees that there are no collisions among data messages and also allows the radio components of each non-cluster head node to be turned off at all times except during their transmit time, thus reducing the energy consumed by the individual sensors[10].

Assuming that all nodes can transmit with enough power to reach the BS, if the distance between any node and the BS is less than the distance between this node and its corresponding cluster head , the node will transmit data directly to the BS. Now, each non-cluster head sends its data during their allocated transmission time to its respective cluster head. The last must keep its receiver on in order to receive all the data from the nodes in the cluster. When all the data is received, the cluster head node performs signal processing functions to compress the data into a single signal. Once this phase is completed, each cluster head can send the aggregated data directly to the BS. In this sub-phase, each non-cluster head can turn off to the sleep mode in order to reduce the consumed energy.

By using our proposed schema based on spectral clustering approach, the total consumed energy of each round is given by

$$E_{round}^{SC\ approach} = L(2N \times E_{elec} + N \times E_{DA} + K \times E_{mp} \times d_{to\ BS}^4 + (\sum_{j=1}^{N-K} d_j^2) \cdot E_{fs})$$

Where $d_j < d_0$.

In a given WSN :

If all non cluster heads operates in a non free space node, the total dissipated energy during a round is determined by:

$$E_{round}^1 = L\ (2N \times E_{elec} + N \times E_{DA} + K \times E_{mp} \times d_{to\ BS}^4 + (\sum_{j=1}^{N-K} d_i^4) \times E_{mp})$$

Where $d_i \geq d_0$

If at least one non cluster head operates in a non free space mode, the total dissipated energy during a round is determined by:

$$E_{round}^2 = L(2N \times E_{elec} + N \times E_{DA} + K \times E_{mp} \times d_{to\ BS}^4 + (\sum_{i=1}^{l, l\neq 0} d_i^4) \cdot E_{mp} + (\sum_{i=1}^{N-K\ l, l\neq 0} d_i^2) \times E_{fs})$$

Where $d_j < d_0$ and $d_i \geq d_0$.

If all non cluster heads operate in a free space mode, the total dissipated energy is given by

$$E_{round}^{SC\ approach} = L(2N \times E_{elec} + N \times E_{DA} + K \times E_{mp} \times d_{to\ BS}^4 + (\sum_{j=1}^{N-K} d_j^2 \cdot E_{fs})$$

Consequently ,

$$E_{round}^{SC\ approach} \leq E_{round}^{2} \leq E_{round}^{1}$$

In the proposed idea total dissipated energy during a round is minimal when the SCRC-WSN algorithm is used. A comparison study is made among SCRC-WSN algorithm, LEACH and LEACH-C about the quality and working condition the SCRC-WSN approach shows the improvement of performance in the form of energy and lifetime gains when compared with other algorithms.

Further works are related to the study of other spectral classification techniques such application might be more efficient. Selecting the one which is powerful is the first step in coming work.

V. CONCLUSION

In this paper, we proposed a new approach to deal with the clustering problem in a given WSN. A new protocol is reviewed which is based on the spectral classification method. This approach is an energy-aware adaptive clustering protocol which employs the graph theory and spectral classification to guarantee robust clustering. The adjacency matrix and Laplacian matrixes of the network graph is used in order to run the SCRC-WSN protocol. This protocol introduced strategies allow which to improve the network performances by saving more energy and extending more efficiency the network lifetime.

REFERENCES

[1] Survey on data collection method on WSNs, Neetu.M.Nair,A.Felix Arpkya Jose,2013

[2] Resource aware Data Fusion method for WSN,#spinger, A.abdelgawad and M.Bayoumi,2012

[3] servey on target tracking methods in WSN,K.Ramya,K.Praveen Kumar,Dr. V.Shrinivas Rao(IJCSES),2012

[4] Key management techniques in WSN Kawla Naji,Shnikat,Ayman,Ahman AlQuadah,2012

[5] Multi path routing technique in WSN,Kewei Sha,Jegnesh Gehlot,Rob ert Greve

[6] Localization technique in WSN,Nabil Ali Alrajeh,Maryam Basir,Bilal Shams,2013

[7] Data aggregation technique in WSN, Nandini.S.Patil,Prof.P.R.Patil ,IEEE,2010

[8] Brahim Elbhiri , Sanaa El Fkihi, "A New Spectral Classification for Robust Clustering in Wireless Sensor Networks", WMNC'2013, 978-1-4673-5616-9/13, 2013 IEEE.

[9] Anjan Das, "A Novel Association Rule Mining Mechanism in Wireless Sensor Networks", 978-1-4577-0748-3/12, 2012 IEEE.

[10] Taherkordi, R. Mohammadi, and F. Eliassen, "A communication efficient distributed clustering algorithm for sensor networks," in Proceedings of the 22nd International Conference on Advanced Information Networking and Applications Workshops/Symposia (AINA'08), pp. 634–638, March 2008.

