International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT217218

Redundant ClusterHead Selection and Quality of Data Reconstruction Through Clustering for Large-Scale WSNs

Edali Jasmine Cleticia I*1, Feiroz Khan T H2, Rosy Jacob3
*1PG Scholar, CSE Department, SRM University, Chennai, TamilNadu, India
2Assistant Professor (OG) CSE Department, SRM University, Chennai, TamilNadu, India
3 Scholar Software Engineering Department, University of Westmister, London, UK

Article Info
Volume 7, Issue 2
Page Number : 62-69

Publication Issue :
March-April-2021

Article History
Accepted : 16 March 2021
Published : 20 March 2021

ABSTRACT

WSN is a large network that consists of a group of spatially distributed sensors nodes. Sensor nodes are partial in power, computational capacities and memory. Sensor nodes compactly installed to monitor physical or environmental conditions, such as temperature, pressure, pollutants. This study examines a Cluster Head Selection (CHs) and takes the CH with the maximum outstanding energy node and less broadcast distance between the CH and BS. It discovered ideal stability between data quality, energy expenditure, and community management ease. The key decision is that the proposal of WSN algorithms must be processing-oriented. i.e., the process of energy on both the Clustering and in-network processing, which ensures both energy efficiency and data quality. Hence, it is more operational to achieve the wireless sensor networks' major loads, which persist the network lifetime.

Keywords: Cluster Head Selection, Clustering, Data Quality, Energy Efficiency, Wireless Sensor Network.

I. INTRODUCTION

Wireless Sensor Network (WSN) is a wireless network consisting of spatially distributed autonomous devices using sensors to monitor physical or eco-friendly situations. A WSN machine incorporated commencing that offers Wi-Fi connectivity stressed international and disbursed nodes. A sensor node is a small device used to build up data from its neighbouring region, perform smooth computations, and interconnect with different sensors or Base Station (BS)[2].

Eventually, it is concluded from the survey that, nonetheless, it is hard to find an extra scalable, power-efficient and strong clustering scheme for data gathering in wireless sensor networks [11]. The cluster head can be fixed or variable; the cluster head also reduces the topology overhead because we consider the cluster head connection only. The cluster heads reduce the power rate by scheduling the cluster tasks; they also perform data aggregation tasks inside the cluster.

It provides efficient cluster head selection within the cluster and does not decrease throughput when it comes to an end and provides a data recovery mechanism [3].Therefore, effectively reducing and balancing the energy consumption foe various nodes and extending the network life is the key to designing a new protocol. We consider these two aspects and propose an algorithm; it shows a good result in the energy consumption of data transmission, network life, and network load [7].

The purpose of this artwork is to study CPPCA except and to determine whether or not or now no longer the format of a clustering set of guidelines attending to the dreams of the in-network processing set of guidelines, infamous and the format of the SODCC set of guidelines, in particular, are beneficial every from the energy and the information high-nice elements of view.

In the following, Section II relates the nation of the artwork of cluster and in-community technique algorithms for WSN. Section III describes the community version and the specific algorithms used in these paintings. The carried out PC simulations and the acquired consequences defined in Section IV, and the acquired consequences are analyzed in Section V. Ultimately, Section VI summarises this work's conclusions.

II. STATE OF THE ART WORK

In-network process algorithms for WSN attempt to make the most of the information's spatial correlation to extend the networks period. Examples are the Approximate Data Collection or Approximate Data Gathering (ADG) approaches [12], [8], wherever representative nodes are elect for every cluster, and better reconstruction error listed for higher energy potency and WSN lifetime.

The reconstruction error mostly restricted prediction techniques, e.g. in Adaptive Sampling Approach (ASAP), wherever the information is extracted from the network victimization model primarily based on Prediction and reducing transmissions.

Artificial models are used to assess energy consumption and data quality, but the cost of this centralized selection of representative nodes is not analyzed. A similar Prediction oriented approach is adopted in [10] but with a less thorough approach in terms of the reconstruction error.

The energy consumption in a WSN minimized by allowing only some nodes to communicate with the base station.These nodes called cluster-heads [4] [6] [9] gather the information despatched by using every node in that cluster, compressing it after which transmitting the aggregated records to the base station. The model is suitable considering the amount of redundancy found in WSNs; in the transmission mechanism of information, the base station can consume a great deal of transmit power from every node.

The version is appropriate thinking about the quantity of redundancy determined in WSNs; within side the transmission mechanism of information, the bottom station can devour an incredible deal of transmitting electricity from each node. In the current work, each the electricity performance and the reconstructed records place unit's fine assessed for numerous cluster and in-community manner algorithms mixtures.With this analysis, we display that an electricity-saving orientated layout for a set of rules does no longer cause common electricity performance as more than one element is now considered a priori act as electricity sinks.

III. SYSTEM MODEL

The number of redundancies positioned in WSNs; direct transmissions to the bottom station will devour a massive quantity of transmitting electricity from each node. In the clustering method, in preference to sensor nodes sending the statistics to the bottom station proper here, information date ship to the cluster-heads. This paper desires to keep the entire community's electricity intake through this structure because the reality node communicates with cluster head & CH further communicates with BS consumes a good deal much less electricity than non - clustering method.The network model used in this work first described. In short, the WSN modelled by a graph, the nodes are clustered using a specific clustering algorithm, and local data processing performed in each cluster. We describe all the cluster and in-network process algorithms utilized in this work in separate subsections.

Hierarchical routing performs better than Flat routing protocol. Hierarchical routing protocols found to be more energy-efficient than all other variants mentioned in Figure 1 for bigger WSNs [1]. It is also known as the cluster-based routing protocol. In this protocol, all the nodes divided into some small groups, known as a cluster. Every cluster has a coordinator or cluster head (CH). All the cluster members send their sensed data to the CH, and CH sends the data to the BS either directly or via other CHs. These protocols help increase the overall network lifetime and scalability of the network. The messages are sent to the BS using single or multi-hop routing, and the number of messages will be less from the flat routing protocol.The Hierarchical routing protocol split into two types routing –dynamic hierarchical routing, wherein clusters fashioned dynamically and static hierarchical routing. As soon as the clusters fashioned, they stay equal all through the community lifetime. Apart from electricity efficiency, the overtly shared broadcast medium that the WSNs use for conversation makes them susceptible to safety threats [5].

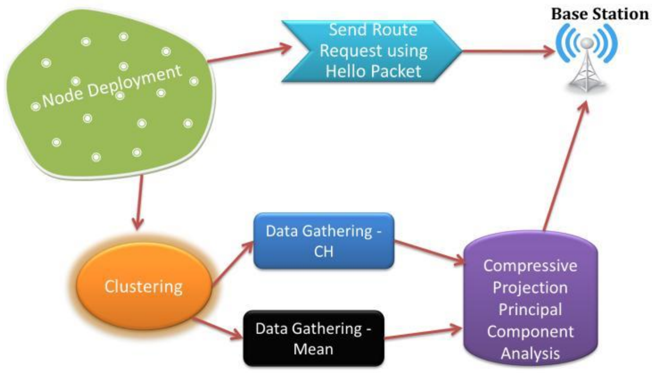

Figure 1. The process of cluster heads selection and data reconstruction using clustering.

A. Network Model

Consider a WSN with N nodes and diverse wireless connection between them. During the WSN operation, every node takes a measuring each T's time instants, getting a total of M information quantities per node.

Regarding the network organization, the WSN is divided into clusters to avoid the capacity limitations of a flat WSN [2]. The trend in this field is that of self-organized algorithms that usually try to cluster the network "on-the-fly" while seeking to reduce at least one of the following limited resources: 1) energy; 2) bandwidth; or 3) computational capabilities.

Also, every cluster of WSN plays facts processing to lessen the range of facts transmissions. This nearby processing can vary from the best possible (i.e., calculating the typical cost of a collection of nodes) to greater state-of-the-art algorithms. (i.e. discovery of local data correlation [1]).The trade-off between the utilization of the process resources and the transmission potency's improvement enclosed within selecting the process rule, as unfortunate associate selection will result in network flooding and high energy consumption.

All sensor hubs are sent arbitrarily in a given region.

- The nodes are constant and homogenous.
- There is an available BS (BS); all the CHS ship their information to the BS.
- All nodes can ship and get hold of information.
- The nodes might not be capable of ship the information to the BS with the unmarried-hop transmission.
- The valid nodes ship and get hold of information reliably.

B. Cluster head Selection

Initialising the network.

1) The base station can get the location of all the sensor nodes in the monitoring area (ID) and the residual energy of the nodes.

2) The monitoring area is divided into some clusters by the Voronoi diagram and the probabilistic perception model proposed. Select network redundant nodes by the probabilistic attenuation algorithm, and these nodes taken as the first kind of hibernation cluster head node.

3) The death of a current cluster head node makes another redundant node active to be the cluster head. If the death node is a current common node, another redundant node ends dormancy to be ordinary.

4) The survival time estimation algorithm can estimate the network average residual energy if the first kind of cluster head nodes all died.

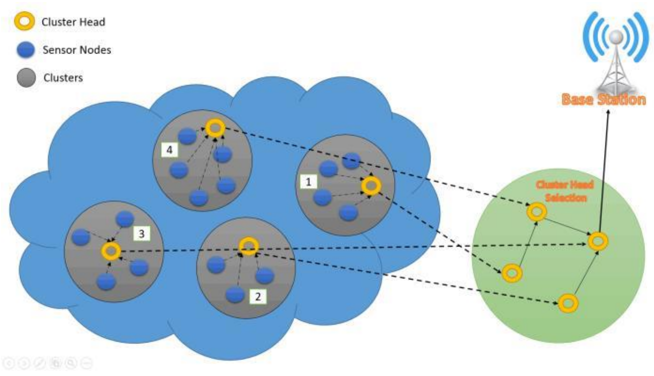

Figure 2. The process of cluster heads selection

Pseudo Code

For Cluster Head Selection
During the selection of cluster head
if
Sensor nodes energy level>= average energy level
then
Sensor node suitable for cluster head
if
Sensor node with highest energy level work as a CH
else
Not eligible for cluster head selection process.

Pseudo Code

For Associate cluster formation
Associate cluster head (ACH) to be selected with the decrease in the energy level of cluster head
if
For the selection of ACH, the node which has higher energy level, after the energy of CH<= Avg. energy act as an ACH.
then
Reduction in energy level for first-round compared to average energy, without selection the CH for the next round, the ACH act as CH for the next round.
else
Resection of CH leads to be a selection overhead.
else if
More energy consumption

C. Approximate Data Gathering – Cluster Head (ADG-CH)

The best ability in-community technique algorithmic application for a WSN and transmits a low-price amount of bytes via the WSN. The key trouble of this set of rules is that, as close by nodes need to have excessive spatial correlation, a unique pattern need to be enough to symbolize all of the measurements made through a collection of close-by nodes. As the

WSN partitioned through a few clustering set of rules, it is miles low-cost to expect that the size from any node that suits into that cluster may want to constitute the measurements from the complete cluster within side the DFC. It is likewise affordable to anticipate that the consultant node ought to be the CH.

A. Approximate Data Gathering – Mean (ADG-M)

It is an improvement of ADG-CH as the bytes transmitted to the DFC are a better representation of the measurements made by a cluster. This algorithm transmits the average value of the cluster's data to the DFC. The disadvantages are that all the nodes have to transmit their measurements to the CH and that the CH spends additional energy computing the measurements' average. On the other hand, the main benefits of ADG-M are that the DFC receives a better approximation of the values measured by the nodes and that the values received by the DFC are not susceptible to errors of the electronics of the CH.

B. Compressive-Projections Principal Component Analysis (CPPCA)

The operation of the algorithm is the following: (i) the measured data by a group of Ni nodes gathered by a representative node (i.e. the CH); (ii) the CH assembles the data matrix X of size Ni × M, and is encoded by performing an orthonormal projection to a lower- dimension random space P of dimension K, with the orthonormal compressed projection matrix P; (iii) the representative nodes transmits the encoded data to the DFC; (iv) the DFC estimates the covariance matrix of the received data (v) the DFC uses the estimated matrix and performs a Projections Onto Convex Sets optimization to resolve the Li principal eigenvectors, and (vi) the DFC uses the eigenvectors to recover the PCA coefficients and obtain the highest possible information about the measured data.

IV. PERFORMANCE EVALUATION

The performance analysis is observed using the network simulator, which most researchers generally use, especially for wireless networks simulation. The simulation parameters used for the communication plans tabulated in Table 1. The network performance metrics like packet delivery, packet loss, throughput and energy are compared against the existing system to get the percentage improvement by using SODCC+CPPCA in WSN.

TABLE 1.

Simulation parameters

Simulation Parameter	Value
Simulation area	800×800m
Number of nodes	50
Simulation Time	50 ms
Initial Energy	1 J
Communication protocol	User Datagram Protocol (UDP)
Traffic Model	Constant Bit Rate (CBR)
Propagation model	Two ray ground
Antenna Type	Omni Antenna
MAC type	IEEE 802.11
Mobility Model	Random waypoint

4.1 Packet Delivery Rate

The Packet Delivery Rate (PDR) is the rate of the entire packets received over the simulation time in the WSN considered. The PDR is calculated by (1).

$$PDR = \frac{Packet\ Received}{Total\ Packet\ Send} \qquad (1)$$

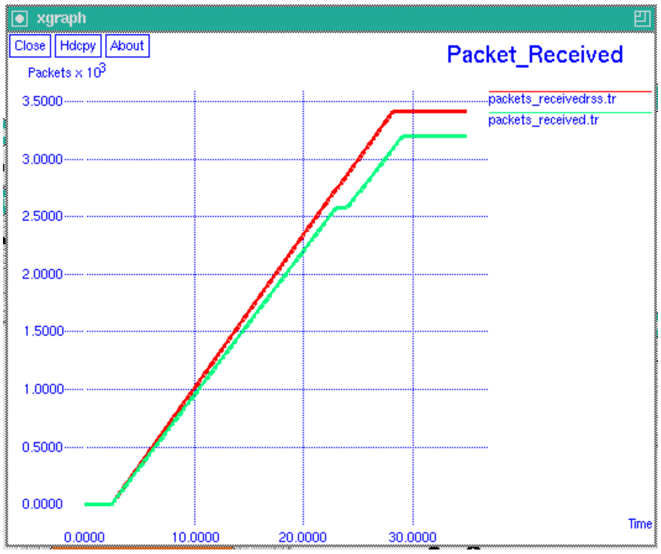

Figure 3. Packet delivery rate.

The PDR of SODCC+CPPCA and LEACH protocol shown in Figure 3. The PDR of the proposed SODCC+CPPCA is greater than the existing LEACH scheme's PDR.

4.2 Average Delay

The average delay is well-defined as the time difference between each packet send and received at any order of time. It is measured by (2).

$$Avg.\ Delay = \frac{\sum_0^n Pkt\ Send\ Time - Pkt\ Recvd\ Time}{Time} \quad (2)$$

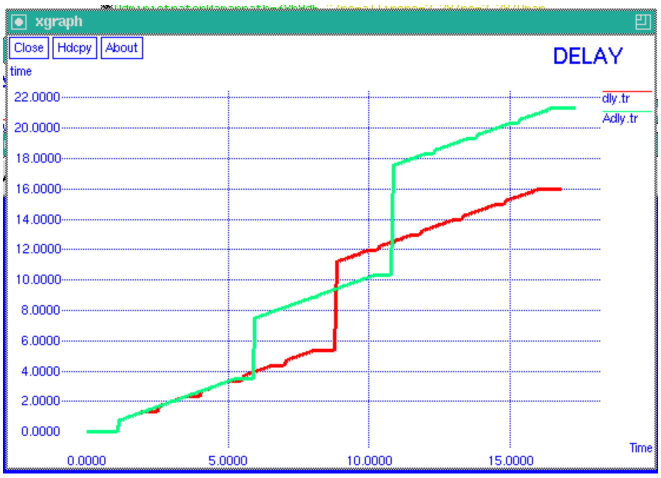

Figure 4. Average delay.

Figure 4 indicates that the delay value is low for the SODCC+CPPCA than the existing scheme LEACH.

4.3 Throughput

It is vital parameters for routing with inside the network. Figure five indicates the throughput at numerous time durations for each LEACH and SODCC+CPPCA protocols. Throughput is estimated by (3).

$$Throughput = \frac{\sum_0^n Pkt\ s\ Received\ (n) * Pkt\ Size}{1000} \quad (3)$$

It can be observed from Figure 7 that the throughput increased in the proposed method than the existing method.

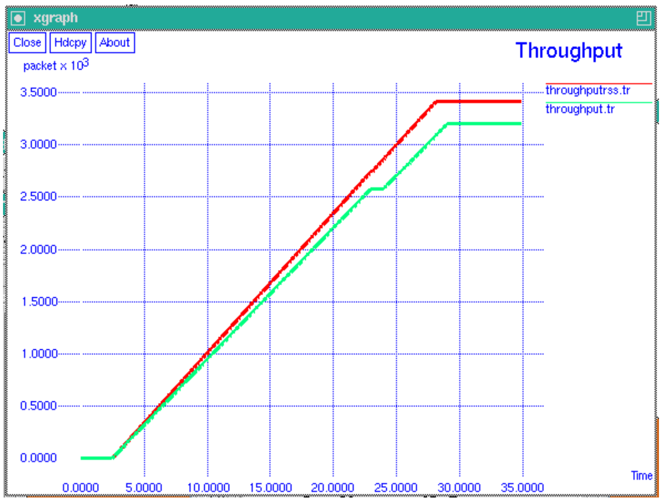

Figure 5. Throughput.

4.4 Residual Energy

A measure of the residual energy gives the rate at which the network operations consume energy. Energy consumption is the typical energy used up by all the nodes in the network. The nodes' initial energy is 1J, and for every operation, there is a periodic reduction of energy for every node.

The difference between the LEACH protocol's average energy consumption and the SODCC+CPPCA protocols shown in Figure 6.

Figure 6. Residual energy

V. CONCLUSION

To increase network lifetime by increasing node lifetime, an energy-efficient routing algorithm along with a novel cluster head selection process proposed here. The CHs are selected based on the parameters such as residual energy of a node, distance of the node from the BS and how recently the node selected as cluster head. The advantage of using a single-hop and multi-hop routing schema is to ensure BS reachability from all network nodes. Performance comparison shows energy consumption in proposed work is less than that if CHs are randomly selected. The community lifetime extended as well. Also, a malicious node barred from its participation as a CH withinside the cluster on this work. As destiny extensions, the adaptive model of the SODCC set of rules permits the reconfiguration of the clusters following the evolution of the traits of the measured records developed. In this case, the strength expenditure may also rely upon the measured field's fee of the variant. Devising robust algorithms to preserve the balanced performance of the SODCC algorithm in rapidly changing environments will also be the focus of future research.

VI. REFERENCES

[1] K. Akkaya, M. Younis, "A Survey on routing protocols for wireless sensor networks, in Ad Hoc Networks, vol. 3, pp. 325-349, 2005.

[2] Amir SepasiZahmati, BahmanAbolhassani, Ali AsgharBeheshtiShirazi, and Ali ShojaeeBakhtiari, "An Energy-Efficient Protocol with Static Clustering for Wireless Sensor Networks." Internation Journal of Electronics, Circuits and Systems, Vol. 1, No. 2, 2007.

[3] Devesh Pratap Singh, R.H. Goudar, Bhasker Pant and SreenivasaRao, "Cluster head selection by randomness with data recovery in WSN," CSIT, Vol. pp. 97-107, June 2014.

[4] Femi A. Aderohunmu Jeremiah D. Deng Martin K. Purvis (Oct. 2009), "Enhancing Clustering in Wireless Sensor Networks with Energy Heterogeneity", the Information Science Discussion Paper Series.

[5] S. Ghidiyal, A. Gupta, M. Vague, A. Semwal, "Analysis of Wireless Sensor Networks: Security, Attacks and Challenges," International Journal of Research in Engineering and Technology (IJRET), pp. 160 – 164, 2014.

[6] Gilgal Lee*, Jonguk Kong*, Minsun Lee*, and OkhwanByeon* (2005), "A Cluster-Based EnergyEfficient Routing Protocol without Location Information for Sensor Networks," International Journal of Information Processing Systems vol. 1, No. 1, pp. 49-53.

[7] Hao Wang and PuTu, "A Redundant and EnergyEfficient Clusterhead selection protocol for Wireless sensor network," Second

International Conference on Communication Software and networks, 2010.
[8] C. C. Hung, W.C. Peng, and W.C. Lee, "Energy-aware set-covering approaches for approximate data collection in wireless sensor networks," IEEE Trans. Knowl. Data Eng., vol. 24, no. 11, pp. 1993-2007, Nov. 2012.
[9] Indranil Gupta (March 2005), "Cluster-head Election using Fuzzy Logic for Wireless Sensor Networks," Master's thesis Dalhousie University Halifax, Nova Scotia.
[10] H. Jiang, S. Jin, and C. Wang, "Prediction or not? An Energy-Efficient framework for clustering-based data collection in wireless sensor networks," IEEE Trans. Parallel Distrib. Syst., vol. 22, no. 6, pp. 1064-1071, Jun. 2011.
[11] K. Ramesh and Dr K. Somasundaram, "Comparative study of cluster head selection algorithms in wireless sensor network," International Journal of Computer Science & Engineering Survey(IJCSES) Vol. 2, No. 4, November 2011.
[12] Wang, H. Ma, Y. He, and S. Xiong, "Adaptive approximate data collection for wireless sensor networks," IEEE Trans. Parallel Distrib. Syst. Vol. 23, no. 6, pp. 1004-1016, Jun. 2012

Cite this article as :

Edali Jasmine Cleticia I, Feiroz Khan T H, Rosy Jacob, "Redundant ClusterHead Selection and Quality of Data Reconstruction Through Clustering for Large-Scale WSNs", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 2, pp. 62-69, March-April 2021.
Available at
doi : https://doi.org/10.32628/CSEIT217218
Journal URL : https://ijsrcseit.com/CSEIT217218