International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 12, No.1 (Sep-2022)

https://dx.doi.org/10.12785/ijcds/120159

Novel Framework for Enhancing Data Quality using Data Correlation Factor in Wireless Sensor Network

Anand Gudnavar[1] and Dr. Manjanaik N[2]

[1] Research Scholar, UBDTCE, Davanagere, India
[2]Professor, Dept. of ECE, UBDTCE, Davanagere, India

Received 07 Mar.2022, Revised 04 Apr. 2022, Accepted 19 Jun. 2022, Published 30 Sep. 2022

Abstract: Sensors are randomly deployed in Wireless Sensor Network (WSN) and it is observed that the sensor density is increasing rapidly. Owing to which there is presence of similar forms of singular event data captured by different sensors. This problem of data correlation is closely associated with cluster formation process, in which energy efficiency is emphasized compared to data quality. Existing reviews show that clustering approaches require major revisions in order to ensure better data quality. Therefore, the proposed system introduces a clustering mechanism that uses data correlation as the essential parameter for the selection of clusterhead, in order to control the transmission of error-prone data packets during the process of data aggregation. Using analytical research methodology, the proposed system introduces three sequential clustering algorithms for ensuring better selection of clusterhead on the basis of best data correlation value. The simulated outcome of the proposed study shows that, it offers better data quality in contrast tothe existing system.

Keywords: Data Quality, Redundancy, Clustering, Selection, Wireless Sensor Network

1. Introduction

The utilization of the Wireless Sensor Network (WSN) has been consistently increasing owing to the beneficial factors associated with remote communication demands [1]. It offers the capability to sense various forms of physical as well as environmental attributes and transmits the collected data to the base station. The data transmission carried out by WSN is called as *data aggregation*where different sensors collect the data and forward the aggregated data to the sink using either single or multiple hop [2] [3]. The process of collecting the data is called as *data fusion* [4]. This entire process of data aggregation is actually initiated using *clustering operation* [5] which arranges the sensors in groups. Each group has one clusterhead and multiple member nodes. The clusterhead in each cluster collects the raw data from member node and performs data fusion in order to remove the redundancies. However, there are multiple problems in this process viz. i) a clusterhead can ascertain non-redundant data from its own cluster but it cannot ascertain if the same data exists within other clusters, ii) usage of single hop-based routing will let forwarding of the fused data from individual clusterhead to sink where there will be no interaction with another clusterhead, thereby causing significant amount of communication overhead over the sink, iii) usage of multihop network will further increase the delay owing to selection of proper routes and hence may affect the data forwarding performance in large network. Therefore, clustering is the most essential part of operation in the data aggregation. According to various literatures e.g.[6] [7], it is said that clustering is important for performing energy-efficient communication in WSN; however there has been no discussion about how clustering also affects data quality, which is about uniqueness in aggregated data from the clusterhead.

Currently there are various algorithms that have claimed to improve clustering operation [8], [9]. These studies do not discuss about data quality or present enough evidence to highlight the influence of it. The research work [10] has associated data quality, with data clustering and data aggregation. It has been observed that existing studies do not have much emphasis on data quality or linked the data quality with data aggregation.

E-mail:anand_gudnavar@yahoo.co.in,manjubdt2009@gmail.com

Thereis good number of possibilities that a same event and its corresponding information are trapped by different sensor nodes residing over different clusters. In such condition, both the clusterhead will treat the obtained information from the member node in a unique manner. An individual clusterhead will assume that it has obtained a unique raw data from the member node, it being unaware of the situation that same unique information is also trapped by other clusterhead. In such condition when they forward the aggregated message then the same message is obtained in the sink and thereby inducing communication overhead. Hence, data correlation is one of the mechanisms to find the degree of the similarity of the data arriving from different sensors. A distributed mechanism of computing data correlation could offer a potential check over the redundancies of data, which could be used for two purposes viz. i) controlling of communication overhead as well as ii) selection of clusterhead on the basis of data correlation factor.

It is to be noted that existing approaches perform selection of clusterhead on the basis of residual energy and distance.These techniques donot control data quality as there is no information, if the neighboring nodes have similar forms of data. Hence, the contribution of the proposed system is to develop a novel data clustering approach that can offer improved significance to data quality in WSN. The proposed system uses data correlational factor for this purpose. The organization of the paper is as follows: Highlights of related work is carried out in Section 2. Briefing of research problem identified after reviewing the existing research is discussed in Section 3 followed by adopted methodology of research is discussed in Section 3.A. Implementation algorithms are presented in Section 4. Result discussion is carried out in Section 5 followed by summary of paper in Section 6.

2. Background

This section discusses the existing research approaches towards data aggregation as a continuation of our prior work [11]. Existing approaches of clustering are more focused on using adaptive methods in order to address energy efficiency [12]. It is also seen that performance of clustering can be significantly enhanced by inclusion of correlation of the data [13]. The data aggregation concept resulted in increased forwarding of massive fused data and hence compression became inevitable. Chidean et.al [14] have used compressive sensing as well as wavelets in order to streamline the data aggregation. The Study towards quality of the sensory data in [15] uses principal component analysis. Harb et al. [16] have presented a rotation strategy of the clusterhead with a target to obtain higher network lifetime. The work in [17] discusses the significance of the special data aggregation technique using statistical approach. Usage of k-means approach improves the performance of energy efficiency in WSN as seen in the work of [18]. Application of time-series is found to improve the process of data aggregation considering single-hop [19]. Usage of learning-based methods claims to offer better clustering performance. This fact is emphasized in [20] where the authors have used self-organizing map in order to improve the clustering performance in WSN. The study outcome showed that self-organizing map offers best performance as compared to other existing methods. The concept of optimization is implemented in [21] where network-based cost factor is highlighted. The said cost factor helps to minimize the communication overhead in WSN. Raja and Datta [22] have introduced privacy-protection for securing data aggregation process. The work of [23] brings-in a unique scheduling scheme towards an efficient power control during clustering operation. Sasirekha and Swamynathan [24] have presented a routing-based approach considering mobile agents for energy efficiency in WSN. Nearly similar direction of the work has been carried out by the authors of [25].They use a typical routing approach to improve the network lifetime of WSN. Usage of compressive sensing is seen in the work of [26] where the signal quality was found to be better. The work of [27] has used tree-based approach for enhancing the query management in dense traffic. Vancin and Erdem [28] have assessed the clustering performance for different energy efficient algorithms in heterogeneous network. It is also noted that there are certain security techniques that emphasize on security aspect of data aggregation methods with better communication performance [29]. Usage of principal component analysis is proven to offer better redundancy management [30].Homomorphic encryption is also found to retain good security balance as well as network balance [31]. Hence, there are various mechanisms to offer improvement in performance of data aggregation in WSN. The next section presents brief discussion of problems that are found unsolved in existing system.

3. Research Problem

The unsolved problems associated with data aggregation are as follows:

- Existing data aggregation mechanism is loosely associated with data correlation as well as data redundancy which results in poor data quality.
- Selection criteria based on similar higher residual energy is less emphasized in clusterhead selection in existing research papers. This is flawed from design viewpoint.
- There is less improvement in the concept of data correlational factor from the viewpoint of large traffic due to which the data quality degrades.
- Currently used techniques lack thecost effective data analysis mechanism while performing data aggregation in WSN.

The statement of the problem associated with review work is "*Developing a computational model for performing clustering operation in order to enhance the data quality using data correlation factor in WSN over challenging environment*."

A. Proposed Solution

The prime aim of the proposed study is to design and develop a novel method that can perform effective data aggregation with the aid of a novel clustering mechanism. The novelty of the proposed system is that it performs dynamic selection of clusterhead on the basis of the data correlational factor. The proposed scheme of implementation is shown in Fig.1.

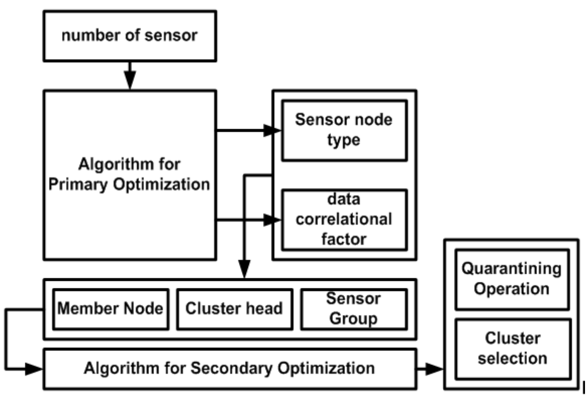

Figure 1. Proposed Flow

The proposed system implements analytical research methodology where a novel clustering method has been introduced in order to retain superior degree of data quality. Unlike the existing systems, the presented system performs the selection of clusterhead on the basis of data correlational factor. The prime concept introduced in our work is that, if data correlation factor of the sensors are computed then, a progressive search happens towards the best correlation factor that leads to selection of best clusterhead with unique and non-repeating data. At the same time, the nodes that are not selected as a clusterhead will be treated as candidate clusterhead and will, still be considered for correlation analysis in order to eliminate any form of uncertainties while computing final correlation value. This phenomenon assists in consideration of any form of dynamicity of the topology in case of variable location of base station. This also assists in confirming nearly similar performance of the clustering, as well as assurance to equivalent data quality irrespective of any position of the base station. The essential contribution of the proposed system is that, it should reduce the flow of the similar data stream which are evaluated on the basis of the data correlational factor. This process offers a good balance between data quality and performs good optimization of the resources resulting in efficient data aggregation.

4. ALGORITHM IMPLEMENTATION

The prime purpose of the proposed system is to introduce a clustering approach that utilizes optimal value of data correlation for carrying out clustering operation. The clustering operation is carried out in such a way thatthe final outcome leads to superior quality of aggregated data over the base station. For this purpose, four discrete algorithms are formulated that leads to better form of data quality. Discussions of these four algorithms are as follows:

A. Algorithm for generation of Proximity Node Buffer

This algorithm is mainly responsible for deploying the sensors in the simulation area in random fashion. However, it also constructs a routing memory that is utilized for evaluating the information of the data associated with data quality. The steps of the algorithm are as follows:

Algorithm for Generation of Proximity Node Buffer
Input: A (simulation area), n (sensors), BS (base station)
Output: rmem(routing memory)
Start
1. init A, n(x,y)
2. A→rand(n, BS)
3. Construct rmem(dval, nad, ntype)
4. dmap→d/argmax(d)
5. For i=1:n
6. ix→f(edis≤Trange)
7. End
8. update rmem
End

The algorithm takes the input of A (simulation area), n (sensors) with unique positions (x, y), and BS (base station), which after execution results in rmem (routing memory). An interesting part of the implementation is that it offers random placement of base station unlikely existing approaches (Line-2). This leads to formation of the routing memory (Line-3) which is characterized by actual value of data *dval*, adjacent nodes *nad*, and type of sensor *ntype*. In line-4, the data *d* is computed by evaluating averaged effective data. The proposed system further performs assessment in order to find the adjacent nodes using the condition shown in Line-6. The algorithm checks for unique distance using function *f(x)* if the Euclidean distance *edis* between two sensors is less than transmission range *Trange* (Line-6). Finally, the routing memory is updated in order to obtain the proximity node buffer information (Line-8).

B. Algorithm for Primary Clustering

This algorithm initiates the clustering operation on the basis of correlation value. The goal of this algorithm is to perform selection of cluster head and distinguish it from non-clusterhead while in the process of initiating clustering operation. The steps of this algorithm are as follows:

```
Algorithm for Primary Clustering
Input: n (sensor),
Output: rmem
Start
1. For i=1:n
2.    Δd<ThresDM
3.    χ++
4. End
5. qcorr→g(n, α, ThresDM, d)
6. If qcorr≠0
7.    ni→rmem
8.    compute intracl, intercl.
9.    update rmem
End
```

The algorithm takes the input of *n* (sensor) which undergo processing of clusterhead selection operation. For all the sensors (Line-1), the algorithm obtains the effective distance *Δd* between individual sensor and proximity nodes, whose information can be obtained from routing memory rmem in first algorithm. It is initially compared with the threshold of data matrix *ThresDM* (Line-2) which ensures that whatever data is considered for the evaluation, should reside within the maximum limit of the memory. The algorithm maintains a counter χ which keeps a track of frequency of the data being populated in the routing memory (Line-3). The algorithm increments χ followed by comparison of α value. This step is focused on retaining the minimum score of the data point as the cut-off. The algorithm applies a function *g(x)* for the computing the quantified correlation using the input arguments of *n* (sensors), α (cut-off), *ThresDM* (threshold for data matrix), and *d* (data) (Line-5). After computing the quantified correlation *qcorr*, the algorithm checks if there is a presence of positive value of *qcorr* (Line-6). For positive case of *qcorr*, the algorithm progressively checks if the effective data distance *Δd* is within the threshold of data matrix. Algorithm uses this information to construct matrices *intracl* and *intercl.* If the data distance is within the threshold limit the shortlisted nodes become a part of same cluster i.e. *intracl* or else it is considered to be a part of different cluster *intercl*. All the shortlisted nodes are retained in a matrix *intracl* and *intercl* (Line-8). While doing this computation, the study considers updating the routing memory simultaneous for all the proximity sensors (Line-6 and Line-9). However, the process flags the node as non-clusterhead when the *qcorr* value is found to zero.

C. Algorithm for Secondary Clustering

This is the continuation of the primary algorithm that is responsible for further confirming the formation of the clustering process. The algorithm assists in the formation of an internal message to be used in the clustering process. The algorithm is also followed by an effective management of the memory system in order to ensure that they have optimal utilization of the memory. Comparison of the quantified correlation is carried out on this perspective to formulate final clusters along with the presence of the unique cluster and non-cluster nodes. The steps of the proposed algorithm are as follows:

```
Algorithm for Secondary Clustering
Input: n (sensor)
Output: groupmem (group memory)
Start
1. If qcorr≠0
2.    nmem→[nmem(intercl, intracl),msg]
3. End
4. For j=1:size(n)
5.    msg(ni)→qcorr(j)
6.    nmem→[nmem(j), msg]
7.    For k=1:length(internode)
8.       If qcorr(internode(k)>qcorr(k)
9.          highCorr→internode(j),qcorr(internode(j))
10. For l=1:size(nodebuffer)
11.      If nodebuffer(l))=1
12.         groupmem=(groupmem(l), nodebuffer(l))
13.      End
14. End
End
```

The algorithm initially constructs a temporary memory system for all the nodes based on which the finalization of the *qcorr* value is carried out. The advantage of this algorithm is that it offers a good balance between the extraction of a unique correlation value as well as good control of memory overhead. If the correlated value is found to be positive (Line-1) than the algorithm considers all the sensors that are part of different clusters i.e. *intercl* as well as all the sensors that are part of same cluster as *intracl*. For all the sensors (Line-4), the algorithm constructs an internal message msg where both the type of nodes is considered. A temporary node memory system *nmem* matrix is constructed (Line-2). The algorithm checks for the size of node buffer by accessing all sensors (Line-4). This step initiates the forwarding of the message of proximity nodes with respect to the *qcorr* value. The algorithm further updates this updated value of the *nmem* and internal message *msg*. The algorithm then considers all the nodes of different clusters and checks if the present correlation value of the sensors is higher than the quantified correlation. The algorithm then computes the highest value of the correlation (Line-8 and Line-9). Similar computation is also carried out for nodes within a cluster also. Finally, the algorithm accesses the routing memory for all the size of the memory of the node (Line-10). If the value of the node buffer is found to be maximum i.e. 1 (using probability), then the group memory matrix is formed which retains information about entire local cluster system with respect to the *nodebuffer* (Line-12). The interesting part of this algorithm is that it is capable of assessing all the sensors followed by the extraction of the discrete and unique group memory, which contributes to further memory saving, as well as faster process of data aggregation as only non-repeating values of data will be stored in the message. Another beneficial part of this algorithm is that it performs spontaneous update of the routing memory thereby making the process of data aggregation faster and unique.

D. Algorithm for Tertiary Clustering

This is the final stage of correlation-based clustering process, which is responsible for finalizing the ultimate selection of clusterhead. The algorithm mainly search for the best clusterhead ID and selects the candidate nodes from group memory from prior algorithm followed by further updated version of the data correlation to confirm its selection of the clusterhead. The significant steps of the proposed algorithm are as follows:

```
Algorithm for Tertiary Clustering
Input: n (sensor)
Output: ch (finally selected clusterhead)
Start
1. For i=1:n
2.    For j=1:length(gindex)
3.       [qcorrhigh, ID]→argmax(tqcorr)
4.       If qcorrhighj>tdatacorrj
5.          thighID→tqcorr()
6.       Else
7.          highID→[highID thighID]
8.    End
9.    If tdatacorr≠highID
10.      highcorr→[highID highqcorr]
11.   End
12.   If qcorr≠0
13.      difference=|d(i)-d(tdatacorr)|
14.      ch→unique(CHID)
15.   End
16. End
End
```

The algorithm takes the input of all prior parameters that after processing leads to confirmation of the clusterhead *ch*. Initially the algorithm constructs a temporary memory for all the sensors (Line-1), followed by computation of highest correlation value of each node and stores it in temporary memory *tdatacorr* (Line-3 and Line-4). If the value of new correlation *tdatacorr* is not found to be equivalent to sensorsthen reformulation of the clustering takes place which further updates the newly constructed memory. This operation assists in highly reduced memory consumption. The algorithm then constructs a matrix *tqcorr* with respect to newly formulated memory where using maximum arguments formulation, highest value of *tqcorr* is obtained and stored in *qcorrhigh* matrix along with the identity of the candidate clusterhead (Line-5). If the value of the highest quantified correlation *qcorrhigh* is greater than currently computed data correlation *tdatacorr* then a temporary candidate clusterhead *thighID* is obtained (Line-5). Otherwise the candidate clusterhead is selected on the basis of *tdatacorr* only (Line-5 and Line-7). If the *tdatacorr* is found not matching with recently selected best candidate node *highID* (Line-9) then highest correlation is computed on the basis of *highID* and highest correlation value *highqcorr* (Line-10). The next part of the algorithm implementation is followed by classification of the clusterhead from candidate clusterhead on the basis of the current correlation value. Therefore, the final steps of implementation of the algorithm checks for all the sensors (Line-12) with a positive quantified correlation *qcorr.* This is followed by computation of difference of the data correlation value in order to obtain updated correlation value. The difference is obtained between current data and recently computed correlated data (Line-13). This step is followed by selection of the unique clusterhead with respect to its identity (Line-14).

One of the interesting parts of the implementation of the tertiary clustering process is that it carries out the final classification of the nodes with unique correlated data. The idea is mainly to perform a final check over the correlated value of the candidate clusterhead with the recently updated correlation value for all the proximity sensors. This operation is highly progressive ensuring the selection process of the clusterhead from many candidate clusterhead with nearly distinct correlated data takes place with lesser delay.

Added to this, the memory usage is negligible as a large number ofintermediary temporary matrices constructed get discarded and their heap memory is released while the correct clusterhead is chosen at the end step.This operation offers higher degree of control over the computational complexity associated with the buffer management that finally results in a unique clusterhead with possession of uniquely aggregated data. Therefore, all the four sequential process of algorithm implementation results in higher quality of data.

5. RESULT ANALYSIS

The analysis of the proposed study towards retaining better performance of data aggregation in WSN is carried out using MATLAB. The implementation of the proposed logic was carried out considering 500-1000 sensors with communication radius of 6 meters. The study considers both the threshold values to be within the range of 0.1-2 i.e. T=0.1-0.3 and α=2. The simulation outcome is shown in Fig.2.

According to Fig.2, it shows that during *initial clustering* the proposed system formulates a set of neighboring nodes among all the deployed nodes that are spatially located from base station. According to the primary optimization execution, the cluster is finalized with a tentative clusterhead as well as member node. The communication between the member node and the clusterhead is initiated during the secondary optimization technique, which is then followed by confirming all the communication links in the form of intra-clustering and inter-clustering operation.This proposed system also performs multi-hop communication system between two different clusters and this operation is continued for all the cycle of data aggregation.

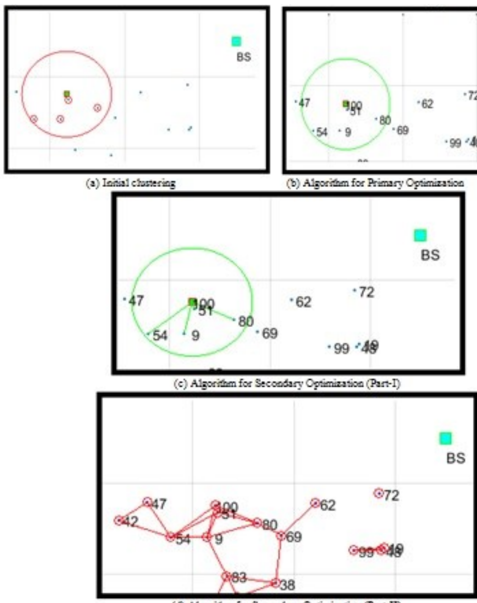

Figure 2. Visual Outcomes of Proposed Simulations

The data correlation factor is ascertained and computed in all the simulation rounds and the node with higher value of data correlation is used for candidate clusterhead, while the finalization is carried out by confirming the clusterhead node (pink nodes in Fig.3) and quarantined node (yellow node in Fig.3). Only the indexer node is considered to be acting as finalclusterhead which significantly reduced the simulation time and increases higher degree of data quality.

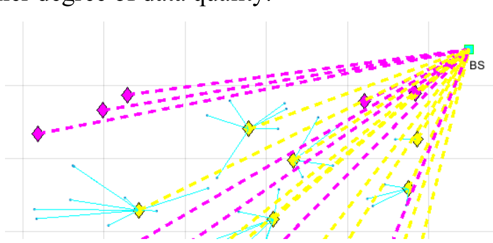

Figure 3. Proposed Data Aggregation

It is clear that proposed study uses the concept of primary and secondary clustering for carrying out unique data aggregation to forward non-redundant data as well as quality-oriented data. The study outcome of proposed system was evaluated with respect to weight-based clustering and hierarchical clustering. The proposed study implements an analytical modeling using simulation-based approach retaining equivalent performance parameter over the comparative analysis. An iterative mechanism was applied in order to assess the degree of consistency. The proposed system performs dispersal of the sensor over the simulation area in a highly random fashion and hence it is imperative to have error performance to have slighter variation. Iterations from 50-3000 have been carried out in order to explore all sorts of changes over the existing hierarchical clustering approach as well as weight-based clustering approach. However, no significant changes were observed after analysis is completed. The alterations of the peaks of curve ofthe proposed system are highly negligible.

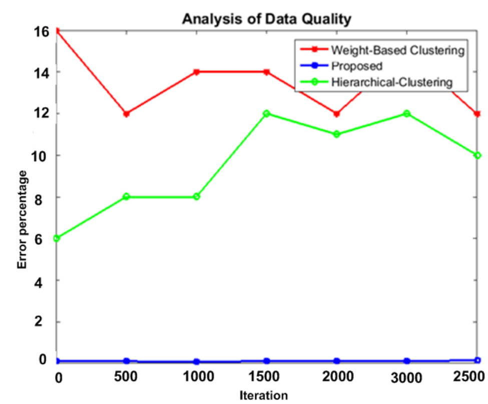

Figure 4. Performance Comparative Analysis of Error

The outcome shown in Fig.4 highlights that the proposed system offers much less degree of variation. The magnified version of the outcome of curve of proposed system was shown in Fig.5 to find that the variation of error is between 0.1-0.140 in max, which is completely within tolerable limits.

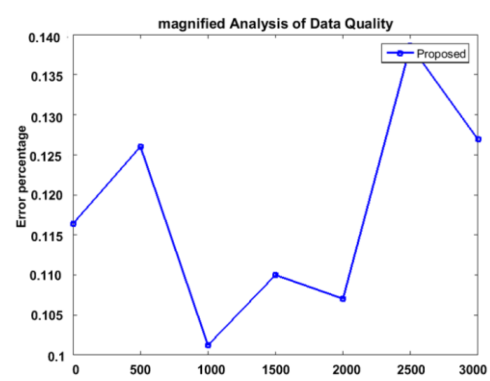

Figure 5. Magnified Visualisation of Error

Added to this it has also been also found thatthe proposed system offers faster processing speed of 0.54322 second as compared to the existing system of average 6.43342 seconds over normal core-i3 processor for 3000 rounds of simulation trails. It is observed that the use of data correlation results in forwarding of error free data packet during data aggregation process effecting in faster operational time with lower memory consumption in a WSN system.

6. Conclusion

This paper presents a discussion of a novel and simple clustering algorithm that performs selection of the clusterhead.Unlike otherconventional clustering mechanism, the proposed system performs selection of the clusterhead on the basis of the data correlational factor where significant reduction in routing time is established due to routing memory. According to this data aggregation process, not all sensors are involved in data forwarding process. In fact, only a set of indexer node is selected that has higher data correlation with its neighboring node and only that node is permitted to forward data on behalf of other non-indexer node. This saves computational resources, retains uniqueness with higher data quality, and offer faster data aggregation process. The study outcome shows that proposed system offers better data quality and faster processing time in contrast to existing clustering technique.

References

[1] Smain Femmam, Building Wireless Sensor Networks: Application to Routing and Data Diffusion, Elsevier, 26-Sep-2017

[2] Fatos Xhafa, Fang-Yie Leu, Li-Ling Hung, Smart Sensors Networks: Communication Technologies and Intelligent Applications, Academic Press, 14-Jun-2017

[3] R Budampati, S Kolavennu, Industrial Wireless Sensor Networks: Monitoring, Control and Automation, Elsevier, 23-Oct-2015

[4] Pramod K. Varshney, Distributed Detection and Data Fusion, Springer Science & Business Media, 06-Dec-2012

[5] Amine Dahane, Nasr-Eddine Berrached, Mobile, Wireless and Sensor Networks: A Clustering Algorithm for Energy Efficiency and Safety, Apple Academic Press, Incorporated, 15-Jun-2018

[6] D. N. Sandeep and V. Kumar, "Review on Clustering, Coverage and Connectivity in Underwater Wireless Sensor Networks: A Communication Techniques Perspective," in IEEE Access, vol. 5, pp. 11176-11199, 2017.

[7] B. Park, J. Nah, JY Choi, IJ Yoon and P Park, "Robust Wireless Sensor and Actuator Networks for Networked Control Systems", Sensors, vol 19, Iss.7, https://doi.org/10.3390/s19071535,2019.

[8] S. Verma, N. Sood and A Sharma "Genetic Algorithm-based Optimized Cluster Head selection for single and multiple data sinks in Heterogeneous Wireless Sensor Network", Applied Soft Computing, vol 85, https://doi.org/10.1016/j.asoc.2019.105788, 2019.

[9] F. Chen, C. Wu, P. Ji and Y. Zhang, "A communication quality improved routing protocol for wireless sensor network," 2009 IEEE International Conference on Automation and Logistics, Shenyang, 2009, pp. 616-620

[10] G. Alvarado, C. Bosquez, F. Palacios and L. Córdoba, "Low-energy Adaptive Clustering Hierarchy protocol and optimal number of cluster head algorithm in a randomized wireless sensor network deployment," 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), Mysuru, 2017, pp. 1-4.

[11] A.Gudnavar,Rajashekar, "Exploring an Effectiveness & Pitfalls of Correlational-based Data Aggregation Approaches in Sensor Network", Springer-International Journal of WIreless Microwave Technology, vol.7, Iss.2, pp.44-56, 2017

[12] I. Atoui, A. Makhoul, S. Tawbe, R. Couturier and A. Hijazi, "Tree-Based Data Aggregation Approach in Periodic Sensor Networks Using Correlation Matrix and Polynomial Regression," 2016 IEEE Intl Conference on Computational Science and Engineering (CSE) and IEEE Intl Conference on Embedded and Ubiquitous Computing (EUC) and 15th Intl Symposium on Distributed Computing and Applications for Business Engineering (DCABES), Paris, 2016, pp. 716-723.

[13] L. Cheng, S. Guo, Y. Wang and Y. Yang, "Lifting Wavelet Compression Based Data Aggregation in Big Data Wireless Sensor Networks," 2016 IEEE 22nd International Conference on Parallel and Distributed Systems (ICPADS), Wuhan, 2016, pp. 561-568.

[14] M. I. Chidean, E. Morgado, M. Sanromán-Junquera, J. Ramiro-Bargueño, J. Ramos and A. J. Caamaño, "Energy Efficiency and Quality of Data Reconstruction Through Data-Coupled Clustering for Self-Organized Large-Scale WSNs," in IEEE Sensors Journal, vol. 16, no. 12, pp. 5010-5020, June15, 2016.

[15] H. Ferng and J. Chuang, "Area-partitioned clustering and cluster head rotation for wireless sensor networks," 2017 International Conference on Machine Learning and Cybernetics (ICMLC), Ningbo, 2017, pp. 593-598.

[16] H. Harb, A. Makhoul, S. Tawbi and R. Couturier, "Comparison of Different Data Aggregation Techniques in Distributed Sensor Networks," in IEEE Access, vol. 5, pp. 4250-4263, 2017.

[17] A. K. Idrees, W. L. Al-Yaseen, M. A. Taam and O. Zahwe, "Distributed Data Aggregation based Modified K-means technique for energy conservation in periodic wireless sensor networks," 2018 IEEE Middle East and North Africa Communications Conference (MENACOMM), Jounieh, 2018, pp. 1-6.

[18] K. Miranda and V. Ramos, "Improving data aggregation in Wireless Sensor Networks with time series estimation," in IEEE Latin America Transactions, vol. 14, no. 5, pp. 2425-2432, May 2016.

[19] B. Nadjib and L. Ammar. "Power-Efficient Routing Based on Ant-Colony-Optimization and LMST for In-Network Data Aggregation in Event-Based Wireless Sensor Networks". International Journal of Computing and Digital Systems,vol 7, Iss 6, pp.321-336, doi: 10.12785/ijcds/070601, 2018

[20] D. Phanish and E. J. Coyle, "Application-Based Optimization of Multi-Level Clustering in Ad Hoc and Sensor Networks," in IEEE Transactions on Wireless Communications, vol. 16, no. 7, pp. 4460-4475, July 2017.

[21] M. Raja , R. Datta, "Efficient aggregation technique for data privacy in wireless sensor networks", IET Networks, 2018

[22] M. Ren, J. Li, L. Guo, X. Li and W. Fan, "Distributed Data Aggregation Scheduling in Multi-Channel and Multi-Power Wireless Sensor Networks," in IEEE Access, vol. 5, pp. 27887-27896, 2017.

[23] H Wala, A.,Jiwa and. &N Alduais. "Data Aggregation Algorithms with Multiple Sensors in Clustered-Based WSN/IoT" International Journal of Computing and Digital Systems, vol 9, Iss 3, doi:10.12785/ijcds/090317, 2020.

[24] S. Singh, S. Chand, R. Kumar, A. Malik and B. Kumar, "NEECP: Novel energy-efficient clustering protocol for prolonging lifetime of WSNs," in IET Wireless Sensor Systems, vol. 6, no. 5, pp. 151-157, 10 2016.

[25] B. Sun, C. Ma, X. Jin and Z. Ma, "Data Aggregation Scheme Based on Compressive Sensing in Wireless Sensor Network," 2016 9th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, 2016, pp. 453-456.

[26] J. Tang, B. Zhang, Y. Zhou and L. Wang, "An Energy-Aware Spatial Index Tree for Multi-Region Attribute Query Aggregation Processing in Wireless Sensor Networks," in IEEE Access, vol. 5, pp. 2080-2095, 2017.

[27] S. Vançin and E. Erdem, "Performance analysis of the energy efficient clustering models in wireless sensor networks," 2017 24th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Batumi, 2017, pp. 247-251.

[28] A. Yessembayev, D. Sarkar and F. Sikder, "Detection of Good and Bad Sensor Nodes in the Presence of Malicious Attacks and Its Application to Data Aggregation," in IEEE Transactions on Signal and Information Processing over Networks, vol. 4, no. 3, pp. 549-563, Sept. 2018.

[29] T. Yu, X. Wang and A. Shami, "Recursive Principal Component Analysis-Based Data Outlier Detection and Sensor Data Aggregation in IoT Systems," in IEEE Internet of Things Journal, vol. 4, no. 6, pp. 2207-2216, Dec. 2017.

[30] X. Zhao, J. Zhu, X. Liang, S. Jiang and Q. Chen, "Lightweight and integrity-protecting oriented data aggregation scheme for wireless sensor networks," in IET Information Security, vol. 11, no. 2, pp. 82-88, 3 2017.

[31] H. A. Babaeer and S. A. Al-Ahmadi, "Efficient and Secure Data Transmission and Sinkhole Detection in a Multi-Clustering Wireless Sensor Network Based on Homomorphic Encryption and Watermarking," in *IEEE Access*, vol. 8, pp. 92098-92109, doi: 10.1109/ACCESS.2020.2994587, 2020.

Anand Gudnavar is working as Assistant Professor, in the Department of Electronics and Communication Engineering at Maratha Mandal College of Engineering Belagavi, India. He received his M.Tech. and B.E. degrees from Visvesvaraya Technological University in 2006 and 2003. He has fifteen years of Academic and Teaching experience. Currently he is pursuing research in the area of WSN data aggregation and is a research scholar at UBDTCE, Davanagere.

Dr. Manjanaik N is currently working as Professor, in the department of Electronics and Communication Engineering at University B D T College of Engineering Davangere, Karnataka. He has more than eighteen years of Research & Teaching experience. He has Obtained B.E degree in ECE at UBDTCE Davangere from Kuvempu University in 2000 and M. Tech degree in Computer Application in Industrial drives from Visevasvaraya technological University, Belagavi, India in 2003 and Completed Ph.D. in Electronics and Communication Engineering from Jain University, Bangalore in 2016. He has published several papers in Journals and Conferences of International repute.