International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-1, November 2019

An Efficient Variable Distance Measure K-Means [VDMKM] Algorithm for Cluster Head Selection in WSN

Ashutosh Kumar Dubey

Abstract**: *Wireless sensor networks (WSNs) provide an empirical and explorative ways to handle and collection of the data in the centralized way. In this paper an efficient variable distance measure k-means [VDMKM] algorithm for cluster heads (CHs) selection in WSN has been presented and analyzed. This approach is divided into two phases. In the first phase data preprocessing has been performed. In this phase size-based weight and threshold assignment have been done on the scaling factor of 1-10. In the second phase VDMKM approach has been applied. The main benefit of this approach is the capability of smaller distance assignments automatically based on different distance measures. The distance measures used here are Euclidean, Pearson coefficient and Manhattan correlation. The complete cluster correlated delay has been calculated along with the packet delivery time. Our distance measure based automatic distance adjustment approach provides better suited distance-based packet delivery in less time. The results are more prominent in the case of Euclidean and Manhattan.***

Keywords**: *WSN, VDMKM, Euclidean, Pearson coefficient and Manhattan correlation.***

I. INTRODUCTION

Wireless sensor networks (WSNs) is a heterogeneous combination of connected sensor. But in terms of energy and bandwidth the resources are very limited [1]. So, the drawbacks which can affect the performances are the limited resources for the communication along with the failure and redundant message receiving problems for the destination nodes [1−6]. Different important factors which are associated with the performances are distance between nodes, the lifetime of the network, energy efficiency, the attack on the carrier node and the channel frequency [7−10]. In this paper an efficient variable distance measure k-means [VDMKM] algorithm has been presented for the automatic node distance calculation with the experimentation to show the variability along with the parametric efficiency consideration.

In 2018, Dahda et al. [11] have been used the combination of neural network, genetic algorithm and dynamic clustering for the purpose of improving the energy efficiency.

Revised Manuscript Received on November 30, 2019.
***** Correspondence Author
Ashutosh Kumar Dubey*, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India.

Their results are proven to be efficient in terms of throughput and reliability. Their results are also found to be better in terms of energy, delay and packet loss. In 2018, Sharma and Sharma [12] have been compared the static and random sink node. It has been followed by the parameters like packet delivery ratio, throughput and end to end delay. Their results show that their approach based on the random sink node is efficient in terms of overall performance. In 2018, Habib et al. [13] discussed about the selection of the most efficient route. They have suggested the complexity during heavy obstacles. They have developed mobile-sink trajectory approach for the WSN which is path-restricted. Their approach is efficient in terms of network lifetime. They have used mixed-integer-linear-program (MILP) model. In 2018, Sandhu et al. [14] discussed about data-driven techniques. They have suggested the parameters like packet delivery ratio, packet drop rate, delay, throughput, and data rate as the performance parameters. They have proposed a framework for data rate prediction which is based on past experience. It has been used for network traversing. Their results support the approach based on the N-fold cross-validation technique. In 2018, Muruganantham and El-Ocla [15] discussed about different routing techniques. They have considered two cases for the WSN implementation. These are Dijkstra algorithm and genetic algorithm testing and then testing and comparison has been performed based on ad hoc on-demand distance vector routing protocol. Their results show that GA has better performance. In 2018, Rubel et al. [16] discussed about the key parameters in WSN like energy efficiency. They have suggested that the critical and non-critical issues should be considered separately. They have suggested clustering for this purpose. They have proposed a scheme based on clustering for the delay-sensitive applications. In 2018, Singh and Kumar [17] discussed the arrangement of nodes through clustering and through the relay nodes. They suggested that the damage nodes may affect this technique. They have focused on the QOS. The discussion is elaborated with theoretical and mathematical aspects. In 2018, Zhang and Fei [18] discussed about the mobile elements. They have suggested the long data delay as the major problem with this approach. They have proposed a data collection algorithm based on mobile elements for this problem. Their results show the capability in the reduction in the data delay. In 2019, Zhou et al. [19] proposed an anonymous routing strategy for preserving location privacy.

Their approach has the capability of setting proxy source node. It has been useful in hiding the location of the real source node. The work of this source node is to select neighbors randomly. Their approach is found to be better in terms of traffic consumption and communication delay reduction. Their approach has also the capability of source node security improvement. There are several research works which has been properly tuned the k-means algorithm with their parameters like [20–22]. The main objective of this paper is to reduce the delay time in the communication by using variable distance measure k-means [VDMKM] algorithm for cluster head selection in WSN.

II. MATERIAL AND METHODS

In this paper an efficient variable distance measure VDMKM algorithm for cluster head selection in WSN has been presented and elaborated. First data preprocessing has been performed. In this step the weight assignments have been performed which is completely based on the size of the packet. The scaling has been considered is 1-10. Random values-based selection and assignments have been performed for all the other scaling except the first. This processing is performed in the source node. Then VDMKM algorithm has been applied in the second phase with different data measures. The distance measures result automatically converted to the threshold values and then the results of the weight segments are automatically arranged according to the distance measures. So, it is useful in automatic selection based on the best suited distance measure. The algorithm of our approach is also shown below to show the approach working phenomena. Figure 1 shows the flowchart of our approach.

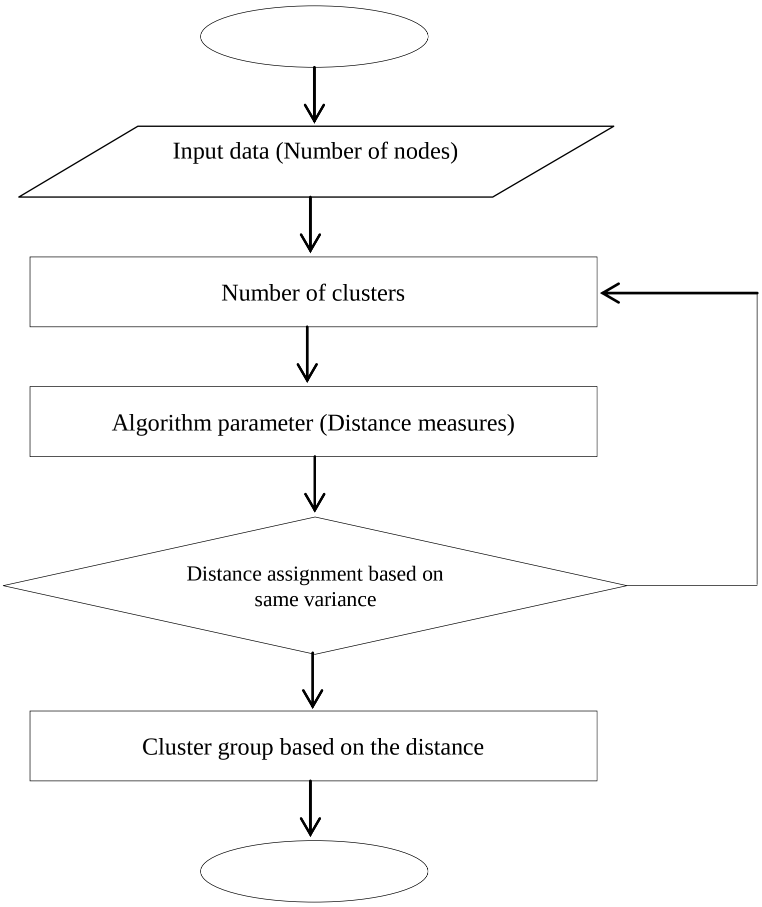

Figure 1 Flowchart for VDMKM

Algorithm: VDMKM algorithm

Step 1: Number of nodes have been initialized.

Step 2: Weight pre-processing has been performed along with the cluster's determination.

Step 3: Iterations have been determined by the same variance along with the random centroid initialization.

Step 4: Distance measure calculation has been performed with the shorter distance allocation

Case 1: Euclidean distance

$d= \sqrt{\sum_{i=1}^{n}(X_i - Y_i)}$

d is the Euclidean distance, X_i and Y_i are the coordinates, n shows the cases numbers

Case 2: Manhattan distance

$d= \sum_{i=1}^{n}|X_i - Y_i|$

Case 3: Pearson Correlation

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sqrt{\sum X^2 - \frac{(\sum X)^2}{n}}\sqrt{\sum Y^2 - \frac{(\sum Y)^2}{n}}}$$

International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-1, November 2019

Step 3: The cluster center is iteratively calculated until the same variance

$c_i = (\frac{1}{n_i}) \sum_{j=1}^{n_i} d_i$

Step 4: Same variance-based automatic assignment between the nodes have been performed and the shorter distance has been picked up.

Step 6: Final selected sets.

III. RESULTS AND DISCUSSION

The results, based on different parameters have been discussed here. For comparative study randomly distributed nodes between 150-300 have been considered. Figure 2 shows the time for automatic distribution of packet delivery based on node assignment in clusters for cycle 1. It clearly shows the time taken is less in all the clustering phases. Figure 3 shows the time for automatic distribution of packet delivery based on node assignment in clusters for cycle 2. Figure 4 shows the time for automatic distribution of packet delivery based on node assignment in clusters for cycle 3. Figure 5 shows the time for automatic distribution of packet delivery based on node assignment in clusters for cycle 4. Figure 6 shows the average time for automatic distribution of packet delivery based on node assignment in clusters. It clearly shows the delay time has been reduced. Figure 7 shows the delay time reduction. Figure 8 shows the average delay time reduction. Figure 9 shows the overall comparative study in packet delivery time. It clearly shows the time taken in terms our approach is less and the approach shows significant improvement in all the intervals. The results are found to be more prominent in the case of Euclidean and Manhattan in comparison to Pearson coefficient.

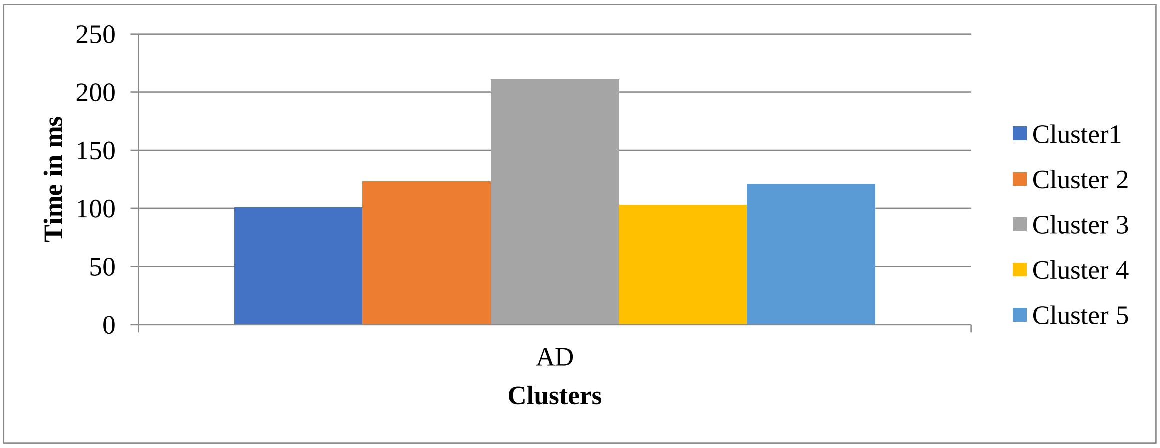

Figure 2 Time for automatic distribution of packet delivery based on node assignment in clusters for cycle 1

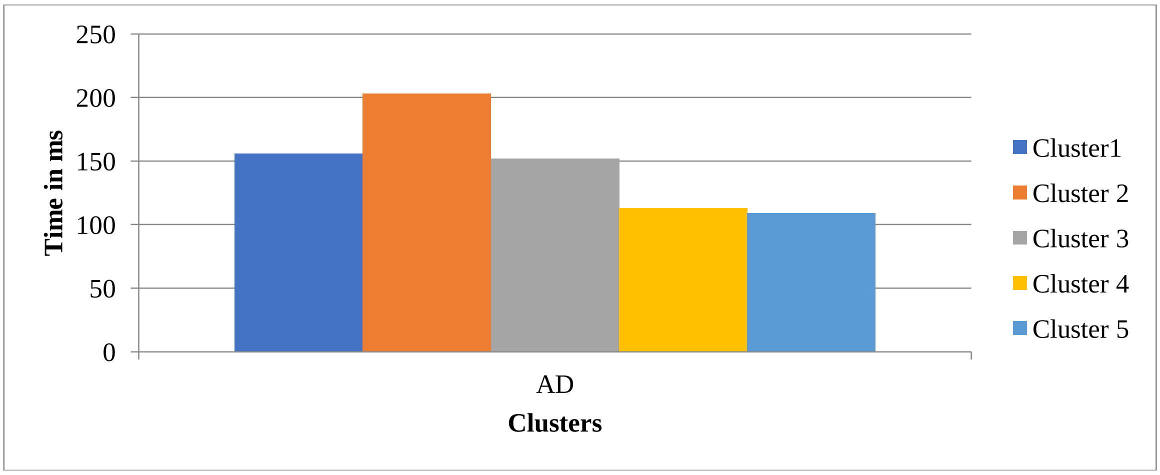

Figure 3 Time for automatic distribution of packet delivery based on node assignment in clusters for cycle 2

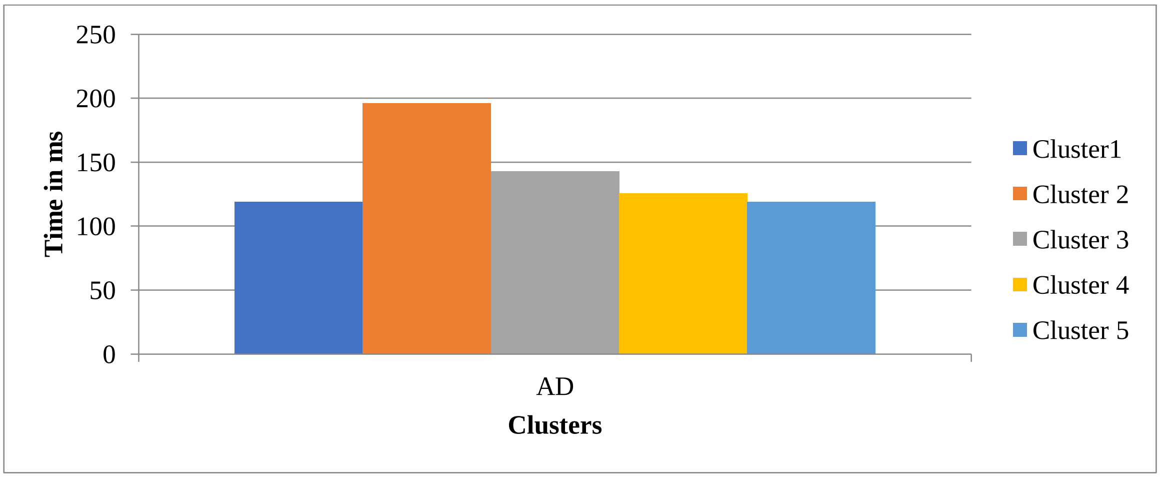

Figure 4 Time for automatic distribution of packet delivery based on node assignment in clusters for cycle 3

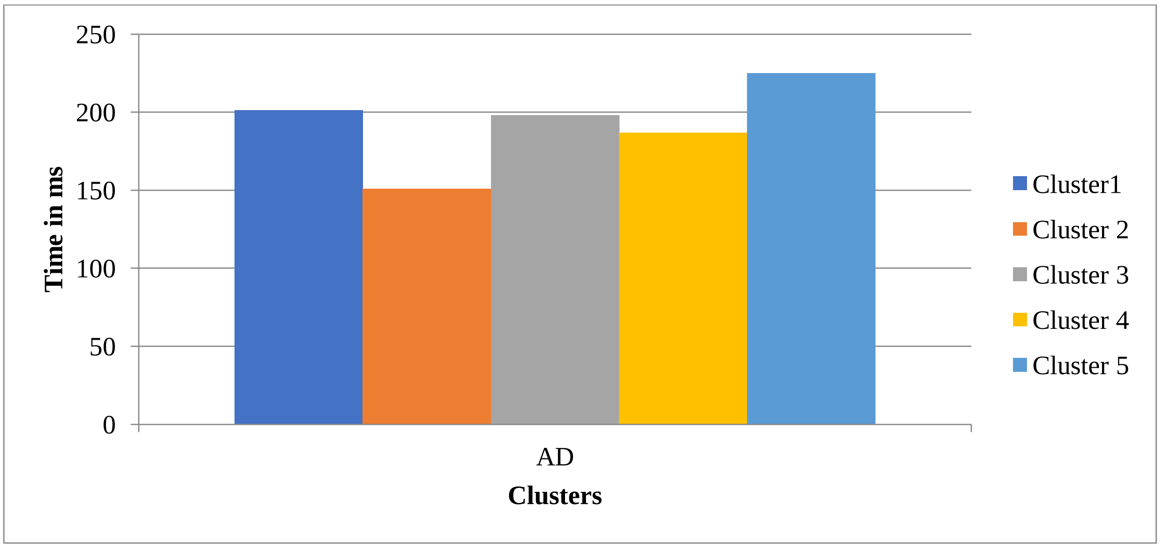

Figure 5 Time for automatic distribution of packet delivery based on node assignment in clusters for cycle 4

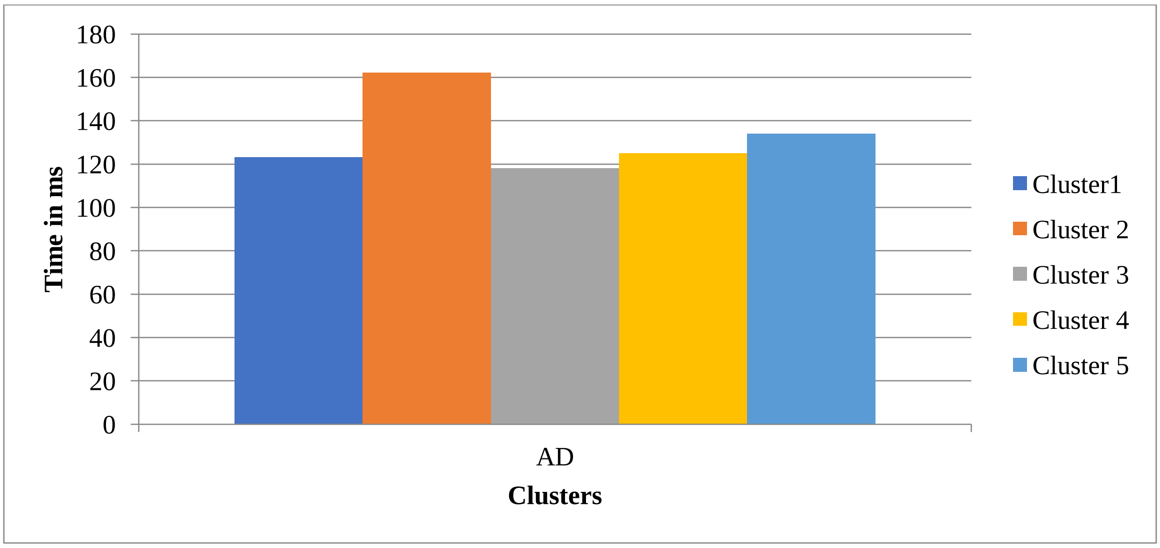

Figure 6 Average time for automatic distribution of packet delivery based on node assignment in clusters

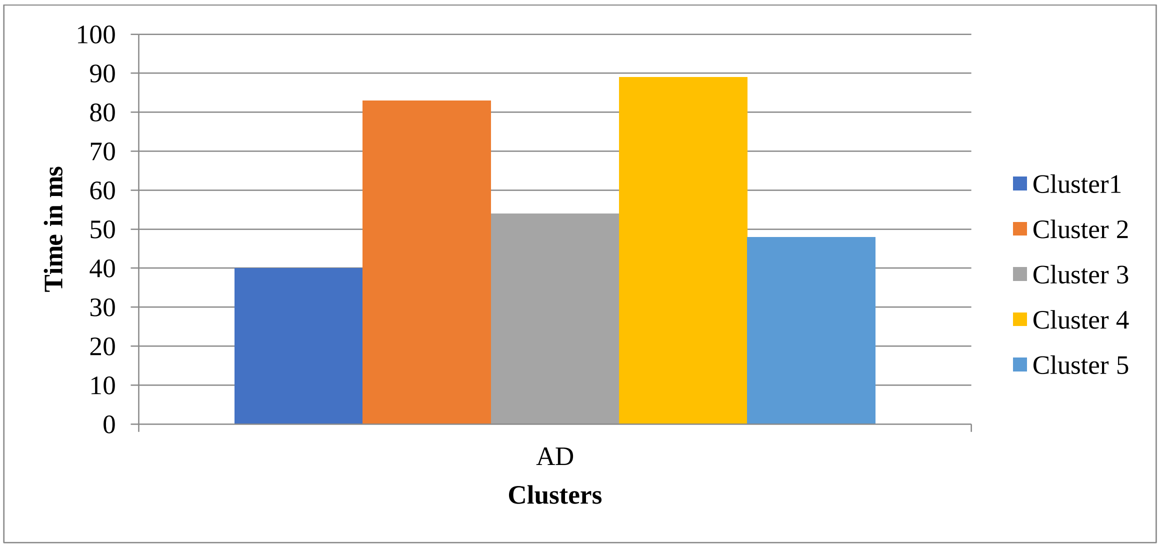

Figure 7 Delay time reduction

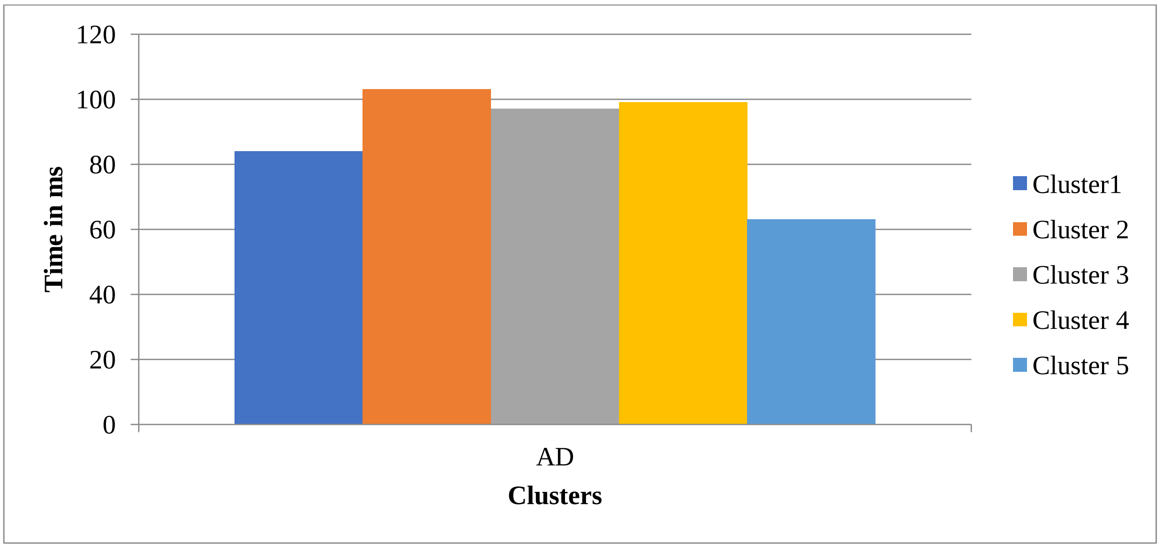

Figure 8 Average delay time reduction

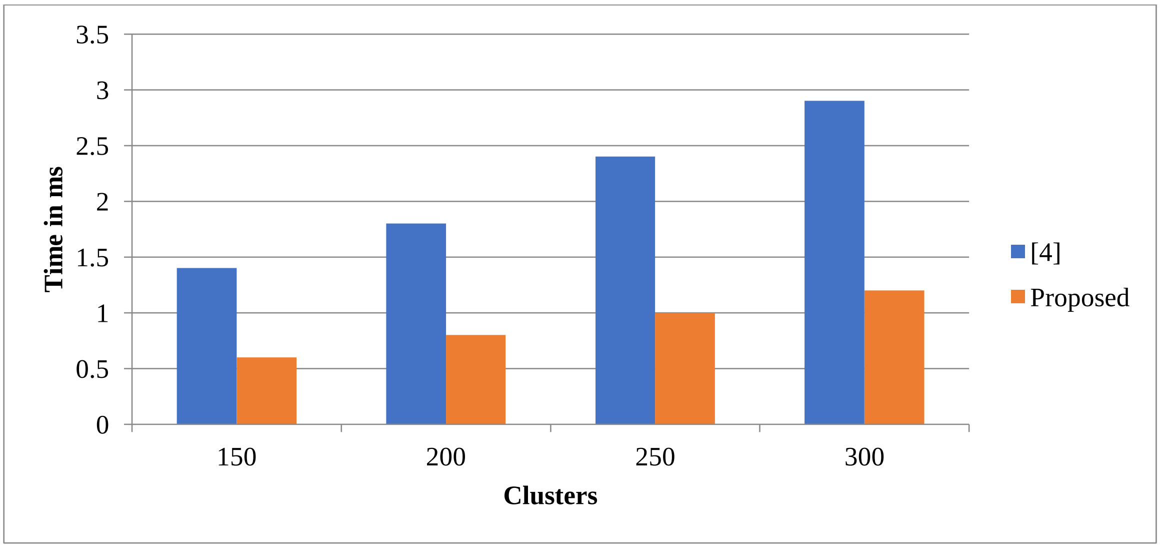

Figure 9 Overall comparative study in packet delivery time

IV. CONCLUSIONS

In this paper an efficient VDMKM algorithm for cluster head selection in WSN has been applied and analyzed. The pre-processing of our approach proved the weight assignment in such manner that it can be scaled in a similar pattern.

Euclidean distance, Manhattan distance and Pearson correlation have been used as the distance measures. The promising results in case of automatic distribution and delay time have been achieved. The results in the case of Euclidean and Manhattan are more prominent. The results in case of packet delivery time are found to be better in terms of related work.

REFERENCES

1. Altoaimy L, Kurdi H, Alromih A, Alomari A, Alrogi E, Ahmed SH. Enhanced distance-based gossip protocols for wireless sensor networks. In annual consumer communications & networking conference 2019 (pp. 1-4). IEEE.
2. Kwiatkowska M, Norman G, Parker D. Analysis of a gossip protocol in PRISM. ACM SIGMETRICS Performance Evaluation Review. 2008;36(3):17-22.
3. Norouzi A, Babamir FS, Zaim AH. A novel energy efficient routing protocol in wireless sensor networks. Wireless Sensor Network. 2011;3(10):341.
4. Khandelwal A, Jain YK. An efficient k-means algorithm for the cluster head selection based on SAW and WPM. International Journal of Advanced Computer Research. 2018;8(37):191-202.
5. Iyengar SS, Brooks RR, editors. Distributed Sensor Networks: Sensor Networking and Applications (Volume Two). CRC press; 2016.
6. Alromih A, Al-Rodhaan M, Tian Y. A Randomized watermarking technique for detecting malicious data injection attacks in heterogeneous wireless sensor networks for internet of things applications. Sensors. 2018; 18(12):4346.
7. Olasupo TO, Olasupo KO. Investigation of factors that determine the performance of wireless sensor systems for large-scale on demand deployment. Universal Journal of Communications and Network. 2017;5(1):1-22.
8. Irandegani M, Bagherizadeh M. Designing an asynchronous multi-channel media access control protocol based on service quality for wireless sensor networks. International Journal of Advanced Computer Research. 2017; 7(32):190.
9. Hammoudeh M, Newman R. Adaptive routing in wireless sensor networks: QoS optimisation for enhanced application performance. Information Fusion. 2015; 22:3-15.
10. Bharti D, Nainta N, Monga H. Performance analysis of wireless sensor networks under adverse scenario of attack. In international conference on signal processing and integrated networks 2019 (pp. 826-828). IEEE.
11. Dahda N, Sindhwani M, Singh C. Energy efficient hybrid technique based on dynamic clustering in wireless sensor network. In international conference on intelligent circuits and systems 2018 (pp. 229-234). IEEE.
12. Sharma HG, Sharma R. A study of static and random sink node in wireless sensor network. In international conference and workshops on recent advances and innovations in engineering 2018 (pp. 1-4). IEEE.
13. Habib A, Saha S, Nur FN, Razzaque A, Mamun-Or-Rashid M. An efficient mobile-sink trajectory to maximize network lifetime in wireless sensor network. In international conference on innovation in engineering and technology 2018 (pp. 1-5). IEEE.
14. Sandhu JK, Verma AK, Rana PS. A data-driven framework for survivable wireless sensor networks. In international conference on contemporary computing 2018 (pp. 1-6). IEEE.
15. Muruganantham N, El-Ocla H. Genetic algorithm-based routing performance enhancement in wireless sensor networks. In international conference on communication and information systems 2018 (pp. 79-82). IEEE.
16. Rubel MS, Kandil N, Hakem N. Priority management with clustering approach in wireless sensor network (WSN). In international conference on digital information, networking, and wireless communications 2018 (pp. 7-11). IEEE.
17. Singh AK, Kumar P. Advancement in quality of services in wireless sensor networks. In international conference on internet of things: smart innovation and usages 2018 (pp. 1-5). IEEE.
18. Zhang C, Fei S. A data collection algorithm with mobile elements in wireless sensor networks. In youth academic annual conference of Chinese association of automation 2018 (pp. 1040-1044). IEEE.
19. Zhou L, Shan Y, Chen X. An anonymous routing scheme for preserving location privacy in wireless sensor networks. In information technology, networking, electronic and automation control conference 2019 (pp. 262-265). IEEE.
20. Dubey AK, Gupta U, Jain S. Analysis of k-means clustering approach on the breast cancer Wisconsin dataset. International Journal of Computer Assisted Radiology and Surgery. 2016;11(11):2033-47.
21. Dubey AK, Gupta U, Jain S. Comparative study of k-means and fuzzy c-means algorithms on the breast cancer data. International Journal on Advanced Science, Engineering and Information Technology. 2018;8(1):18-29.
22. Fakhet W, El Khediri S, Dallali A, Kachouri A. New k-means algorithm for clustering in wireless sensor networks. In international conference on internet of things, embedded systems and communications 2017 (pp. 67-71). IEEE.

AUTHOR PROFILE

Dr. Ashutosh Kumar Dubey received his PhD degree in Computer Science and Engineering from JK Lakshmipat University, Jaipur, Rajasthan, India. He is currently in the department of Computer Science and Engineering, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India. He is also the Senior Member of IEEE and Professional Member of ACM. He has more than 11 years of teaching experience. He has authored a book name Database Management Concepts. He has been associated with many international and national conferences as the Technical Program Committee member. He is also associated as the Editorial Board Member/ Reviewer of many peer-reviewed journals including Elsevier, Springer, BMJ, IOS Press, Bentham Science, Thieme Publishing Group, etc. His research areas are Data Mining, Optimization, Machine Learning, Cloud Computing, Artificial Intelligence and Object-Oriented Programming.