Data Aggregation and Its Impact on Performance Enhancement

Lakshmi Bhaskar, Yamuna Devi C. R.

Abstract: Wireless sensor network incorporates an innovative aspect called as data handling technologies for big data organization. In today's research the data aggregation occupies an important position and its emerging rapidly. Data aggregation incudes, process of accumulating the data at node, then either store or transfer further to reach out the destination. This survey depicts about the previous work on data aggregation in WSN and also its impact on the different services. There are number of data aggregation techniques available for reducing the data, processing the data and storing the data. Some of them are discussed here as a review. The data aggregation performed using certain techniques can also be aimed in having energy efficiency, time efficient, security could be in the form of confidentiality, unimpaired, authenticate, freshness, quality, data availability, access control, nonrepudiation, secrecy, secrecy. These are the relevant performance metrics to maintain the better Qos in WSNs applications. The goal of this paper is to display an overview of existing techniques for performance improvement in homogenous/ heterogenous networks.

Keywords: Data aggregation, Energy efficient, QoS, Wireless sensor networks.

I. INTRODUCTION

Wireless sensor network (WSN) is described as a group of nodes dispersed and separated spatially which are dedicated for the purpose of monitoring and recording the physical environmental conditions and organizing the collected data at a particular location. Latest advances in WSN has exerted numerous new conventions particularly intended for sensor networks where energy mindfulness is a critical concern. These routing conventions may vary from relying upon the application and the network design. To expand the lifetime of Wireless sensor network (WSN), an energy effective plan can be composed and created by means of a calculation to give sensible energy utilization and network for WSN. In order to maintain effective adaptability and quality data aggregation, sensor nodes are regularly gathered into discontinuous, non-covering subsets named as clusters. Clusters construct hierarchical WSNs which fuse effective usage of constrained assets of sensor nodes to effectively reduce energy utilization, so as to support and expand the lifetime of WSN. The protocols discussed here are

considered for multi-hop networks. The paradigm shift from just designing a network to self-learning of the network for dynamically changing the parameter is the growth that is seen as of today.

II. DATA AGGREGATION AND ITS FEATURES

A brief summary of existing set of data aggregation processes available in WSNs and the outcome of each method i.e. performance of these processes is illustrated through this paper, which gives a clarity on different metrics, hence based in them the method can be chosen suitable and proceed ahead in any applications.

Consider an application of WSN such as in Big data, the aspects here are velocity, variety, veracity and value of information to be processed. In order to maintain the efficiency certain solutions needs to be obtained. The WSN has to be supported with effective algorithm for data gathering, analysis and storage. This process can be termed as data aggregation. The methods can be energy efficient, secure etc. for this certain strategy can be considered cluster based, multi-layered aggregation, scalable privacy, fusion based, etc. The combination of different aggregation techniques supports in improvement in performance metrics is observed.

A. Author Boubiche et. al. discusses the new paradigm related to big data analytics using Map Reduce and Hadoop concepts based on their advantages. In WSNs the main challenges are:

- Processing including storage, dimensions, collection, analysis, aggregation and reliability.
- Securing including confidentiality, integrity, authentication, freshness.
- Energy saving.

These challenges are clustered and then solution is designed as they are interlinked. The strategies considered here are not particularly for specific network. Some of the strategies are:

- A distributed algorithm for aggregation based on compression sampling focused on local minimization ,with reduced traffic for data routing in order to dynamically build a routing path.
- For collection of big data through the sensor nodes a multi-layer aggregation infrastructure as well as a priority-based Dynamic Data Aggregation Protocol (PDDA) is implemented.
- In order to recover the original data along with increased accuracy, an effective data reduction method or data compression algorithm for a large scale WSN, on a distributed high-speed wavelet is required.

- A Principal Component Based Data Aggregation (PCA) algorithm supports an efficient transmission of huge amount data sensed with low latency, it also eliminates duplication of data in cluster head nodes to reduce the transmission data complexity.
- A new techniques of data fusion for a wireless multi sensor system, based on a hybrid algorithm for clustering and cluster member selection is very much needed.

As the heterogenous networks are concerned the novel methods need to be developed.[1].

B. Author Dr. M. Preetha   suggests a method using the Energy based Low Energy Adaptive Clustering Hierarchy (E-LEACH) protocol to cluster the sensor nodes into groups. The cluster head are sent with the sensed data from the cluster nodes for further processing, selected based on strategy of highest residual energy among the sensor nodes and its role is to aggregate and compress the data received from all nodes then send to the nearest base station. In this particular algorithm energy consumption is evenly maintained in the network by random rotation of CH's. this method concludes with less energy consumption, lesser transmission cost, network life time and faster simulation time.[2]

C. In order to cater the requirements of heterogenous networks and improved performance the Author Garima through the work proposes a comparison on different reactive heterogenous protocols for performance measurement based on stability period, half dead node and network lifetime. The tool used to analyse is MATLAB. The network lifetime of DRESEP is high due to its energy efficiency. As a future work its suggested that energy efficient selection of circular radius, sink mobility can be used to enhance data throughput in the network.[3]

D. Author Johnu elma in [7] also adds their work related to heterogenous network where the performance improvement is in terms of network lifetime. The methodology employed here is to consider a randomly distributed network and apply EHC algorithm using NS2 simulator. The analysis shows that it has better energy efficiency, faster route identification and packet delivery.

E. There are different ways to have better performance metrics in WSNs. Some of the proposed analysis as discussed below as a brief review:

Author in paper [4] proposes a method to have energy efficient and increased network lifetime for a given network by performing compressive sensing at the nodes and differential optimizations. The method was verified using MATLAB 2010 and compared with other techniques.

As a survey in [5] propose that, its builds an IF algorithm with certain improvements, furnishing the preliminary estimate it can be achieved estimate of the credibility to sensor nodes. A more robust and accurate IF algorithm is designed. Also considers the fact that the aggregator node could be attacked by an adversary. The method describes a secured data aggregation approach over compromised aggregator node.

One of the performance metrics is resilience to data damage due to attack [6]. A spatio-temporal correlation algorithm is discussed it combines the distance from centroid and similarity in order to evaluate degree of attack on each cluster node's aware data, and it could improve precision through weighted calculation for data recovery. Through this method it can also obtain the estimated data value of samples of all clusters through temporal correlation characteristic of the nodes' aware data at different instants of time. For effective measurement of the tampered data it uses the chi-square fitting.

In traditional algorithms the improvements considered are enhanced by the use of genetic algorithms as mentioned in [8]. The purpose is to have node localization and route planning in applications such as Internet of things and internet of vehicles.

The optimization of scheduling a time slot is considered for WSNs [9]. The transmission process not only orients towards reduction in energy consumption through ordinary objectives but also considers method of aggregating the data through a network effectively .

(ii) The global scheduling optimization problem can be changed into distributed as well as collaborative tasks on sensors, because of this feature the proposed approach can be easily accommodated for network environment with dynamic conditions of WSNs, also the computation capability of the sensor device must be exploited.

(iii) By utilizing the local knowledge from the neighbour environment, Q-learning is employed in discovering a process for optimal scheduling. The process of selecting the slot based on time is formulated as a limited state machine.

The data aggregation concept is a method of time slot scheduling in a network. Along with transmission delay, consumption of energy also becomes a relevant measure for its performance.

Although through simulation it can evaluate the theoretical performance of the any proposed scheduling policy in different application scenarios, the results from simulation are not the same since the measurements are from the real time systems in certain cases. Hence, further step is to implement optimization method on the real sensor nodes, then observe and analyse the actual performance on those real time devices. The number of upstream nodes are closely related to the routing structure, it is a crucial parameter in the learning process of the slot scheduling policy. If there is any change in routing structure, it may directly impact the scheduling results. Further some specific techniques should be developed in order to make the learning process simpler, which can adapt to the dynamic changes of routing structure.

F. Some concepts related to reinforcement algorithms [10] are based on Q learning. This process of node deployment is easy and improves network life time. The challenges encountered are:

- The message exchange overhead.
- Routing in WSN using Reinforcement Learning.
- Security concerns of Reinforcement Learning.
- Agent Learning cost.
- Robustness of Reinforcement Learning.
- Energy Consumption.
- Balance between Convergence and Power.
- Efficient Self Learning Algorithm for WSN.

III. RESULTS AND DISCUSSIONS

Yet another aspect of WSN is reducing data redundancy, the algorithm contributing for improving network life time is FTDA it's based on data fusion, i.e. redundancy elimination in the given data and then accumulate, this is fundamental way of reducing energy consumption. The results are based on the five data prediction methods, which are ARMA, AR, MA, non-linear regression and exponential smoothing, in this work it simulates the ammonia nitrogen content data of the National Ecosystem Observational Network. The comparative analysis of accuracy in prediction for five time series methods by experiments is performed. The selection of the an effective exponential smoothing method is done, and also construction of the FTDA data fusion algorithm in a WSN system. The simulation results proves that the FTDA algorithm makes certain improvements in the life time and energy consumption in a network life cycle.

Table- I: Comparison of different data aggregation strategies vs Challenges

Challenges Data aggregation strategies	Energy efficient	Bandwidth Utilization	Clustering	Securing	Network lifetime
Redundancy Elimination Data Aggregation algorithm [12]	Yes	Better than ESPDA	Yes	No	No
Secure Reference based Data Aggregation Protocol [1]	Yes	No	Yes	Yes	No
Energy Efficient and Secure Pattern based Data Aggregation (ESPDA) [6]	Yes	Close to 0	Yes	Yes	No
Secure Data Aggregation Watermarking [1]	Yes	No	No	Yes	No
Distributed compressive data aggregation in Large-scale wireless sensor networks [1]	NO	No	Yes	No	Yes
Sensor data aggregation in a multi-layer big Data framework [1]	Yes	No	Yes	No	No
Lifting wavelet compression-based data Aggregation in big data wireless Sensor networks [1]	Yes	No	Yes	No	Yes
Data aggregation with principal component Analysis in big data wireless sensor networks [1]	Yes	No	Yes	No	No
Scalable privacy-preserving big data Aggregation mechanism [1]	No	No	Yes	Yes	No
A cluster-based data fusion technique to Analyze big data in wireless multi-sensor system [14]	Yes	No	Yes	No	Yes

Another way of data redundancy elimination is proposed in [12], REDA (redundancy elimination data aggregation algorithm) is proposed to reduce energy consumption and also efficient bandwidth occupancy. The study proves up to 44% saving can be done in comparison with ESPDA and

SPDA protocols. This work can be applied for reducing the redundancy in agriculturally based data set and also extended to large area of networks.

The major process in network is clustering, there are different algorithms discussed [13] based on this. It can be hierarchical based or fuzzy based. In this it shows that cluster head selection can be based on different performance parameters like residual energy, distance, centrality of the nodes.

Under adverse conditions the WSN must exhibit an outstanding energy proficiency. To expand the lifetime of Wireless sensor network (WSN), an energy effective plan can be composed and created by means of a calculation to give sensible energy utilization and network for WSN. Clusters make hierarchical WSNs which fuse effective usage of constrained objects of sensor nodes to minimize the energy utilization, this improves the lifetime of WSN. The goal is to display a cutting-edge overview and order of energy effective plans for WSNs [14].

IV. CONCLUSION

There are different methods used for homogenous and heterogenous networks for preforming data aggregation, clustering and routing. During this process the care has to be taken to maintain the security for data, fast routing, energy efficiency, bandwidth occupancy, and network life time. This work shows that there are different techniques available as performance enhancers for an effective network system. Thus, this work can be referred as highlights latest techniques available to us for the purpose of research work and also few applications also. The table 1 shows the comparative study of some of the algorithms used in research recently.

The future work can be carried on improving network lifetime, non-negligible aggregation latency should be taken into consideration during the design process, improvement by energy efficient selection of circular radius. Moreover, the sink mobility can be introduced to enhance data throughput in the network, large scale area considering network complexity in algorithms such as pattern code generation, performing node localizations using genetic algorithms in the applications such as IoT/IoV etc. or from a distributed network. Implementing self-learning features in real time applications by scheduling policy and dynamic changes in routing structure with the help of Q learning.

REFERENCES

1. Djallel Eddine Boubiche et. al. "Big Data Challenges and Data Aggregation Strategies in Wireless Sensor Networks" in IEEE Access Special Section on Real-Time Edge Analytics for Big Data in Internet of Things, Volume 6, 2018, pp.2169-3536
2. Dr. M. Preetha, K. Sivakumar "An Energy Efficient Sleep Scheduling Protocol for Data Aggregation in WSN" in TAGA JOURNAL VOL. 14, ISSN: 1748-0345
3. Garima Singh, Rakesh Kumar "Performance Evaluation of Routing Protocols in Heterogeneous WSN by using Performance Metrics" in International Journal on Future Revolution in Computer Science & Communication Engineering ISSN: 2454-4248 Volume: 4 Issue: 3 March 2018 pp.134 – 138
4. Varinder Pal Singh, Varinder Jit Kaur and Dr. Naveen Dhillon "Performance Evaluation of Compressive Sensing and Differential Evolution Technique based on Energy Efficient Protocol in WSN" in International Journal of Emerging Technologies in Engineering Research (IJETER) Volume 6, Issue 3, March (2018)
5. Madhav Ingle, P.V.R.D. Prasada Rao "Improving IF Algorithm for Data Aggregation Techniques in Wireless Sensor Networks" International Journal of Electrical and Computer Engineering (IJECE) Vol. 8, No. 6, December 2018, pp. 5162~5168
6. Yong Lu et. al. "A resilient data aggregation method based on spatio-temporal correlation for wireless sensor networks" in Lu and Sun EURASIP Journal on Wireless Communications and Networking (2018).
7. K. Johny Elma and Dr. S. Meenakshi "Energy Efficient Clustering for Lifetime Maximization and Routing in WSN" in International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 1 (2018) pp. 337-343
8. Xiao yang Liu et.al "Nonlinear Modelling and Analysis of WSN Node Localization Method" in International Journal of Wireless & Mobile Networks (IJWMN) Vol. 10, No. 1, February 2018.
9. Yao Lu et. al. "Self-Learning-Based Data Aggregation Scheduling Policy in Wireless Sensor Networks" Hindawi Journal of Sensors Volume 2018, Article ID 9647593, 12 pages.
10. Santosh Soni1, Dr Manish Shivastava, "A Review of RL algorithms & techniques for WSN implementation in large scale" in International Journal of Advanced in Management, Technology and Engineering Sciences Volume 8, Issue III, MARCH/2018.
11. Mingxin Yang "Data Aggregation Algorithm for Wireless Sensor Network Based on Time Prediction" in IEEE 2017 978-1-5090-5363-6
12. Sabrine Khriji et.al "Redundancy Elimination for Data Aggregation in Wireless Sensor Networks" in 15th International Multi-Conference on Systems, Signals & Devices (SSD), December 2018.
13. Mrs. Nirupama, Er. Tanika Thakur "Different Cluster Head Selection Approaches of Energy Efficient Protocols in WSN" International Journal of Engineering Trends and Applications (IJETA) – Volume 5 Issue 1, Jan-Feb 2018.
14. Sona Malhotra, Urmila, "Multihop Energy Enhancement Protocols in WSN" in Journal of Network Communications and Emerging Technologies (JNCET) www.jncet.org Volume 8, Issue 2, February (2018).

AUTHORS PROFILE

Lakshmi Bhaskar, received her B. E degree in Telecommunication and 2nd Rank in M. Tech degree in Digital Communication and Networking from Visvesvaraya Technological University, Belagavi. She is currently pursuing Ph.D. in Visvesvaraya Technological University, Belagavi. Her research interests are in Wireless and ad-hoc sensor networks.

Yamuna Devi C. R., is currently working as Head of Department in the Department of Telecommunication Engineering, Dr. Ambedkar Institute of Technology, Bangalore, India. She obtained both her Bachelor of Engineering degree in Computer Science and Engineering branch and PhD in Wireless Sensor Networks from Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bangalore. She has 27 years of teaching experience. Her research interests are Wireless Sensor Networks, Adhoc networks and Internet of Things.