Turkish Journal of Computer and Mathematics Education Vol.12 No.12 (2021), 3390-3395

Research Article

An overview on data aggregation in IOT For wireless sensor network

MOHAMMED SIRAJUDHEEN[1], **DR. A. SHOBHA REKH**[2]

[1]PART TIME RESEARCH SCHOLAR, KARUNYA INSTITUTE OF TECHNOLOGY & SCIENCES, COIMBATORE, TAMILNADU.
[2]PROFESSOR-ECE, KARUNYA INSTITUTE OF TECHNOLOGY & SCIENCES, COIMBATORE, TAMILNADU.

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 23 May 2021

Abstract—in recent days, there is a increasing attention in employing wireless sensor network (WSN) technologies in different scenarios of Internet of things (IoT). On considering the massive smart objects growth with their application, there is a demand to analyze and collect their product data are appropriate of the major challenge. Data aggregation is one of the prominent approaches in the removal of data redundancy and energy efficiency improvement; in addition it enhances the Wireless Sensor Networks lifespan. As well, the competent protocol on data aggregation might decrease network traffic. As a detailed intention takes place in a detailed region, it may be recognized through too much sensor. Considering the main aspects and challenges of data aggregation in WSN, an assessment on various kinds of data aggregation approach and protocol were presented in this approach. The eventual intention of this analysis is to create the basic details to expand novel higher designs depending on the techniques of data integration and clustering which have been presented so far. In this manuscript, various approaches has been described for the reduction of data redundancy and specifically in the aggregation of data.

Index Terms—Internet Internet of things, Data aggregation, wireless sensor network, Networks lifespan, data redundancy.

I. INTRODUCTION

The utilization of sensor network referred as WSN i.e., Wireless Sensor Networks that are spatially scattered autonomous devices to monitor the environmental conditions at different locations like vibration, noise, warmth, motion pollutants, and force [1]. WSN have the multi functioning sensor nodes that include thousands or more number of less power that works in an unattended human area with the limited computational and also have the capabilities of sensing. Also, every node equipped with a small microcontroller, radio transceiver or with some other wireless communication mechanism, and the source like as battery. For monitoring purposes, these less cost, power-efficient sensor nodes works together to built a network. The sensor nodes cooperation that provides the collection and transmission of wide range of message in WSNs that have the monitoring framework such as Temperature, Humidity, etc. to the sink node like as BS (Base Station) that enables the data to the respective user. The Internet of Things (IoT) technology is used in various fields for the purpose of enhancing the performance, but nowadays, the agriculture field also utilizing the IoT for makes the farming smart. The recent survey stated that the world population could reach around nine by the year of two thousand and fifty. Hence there is a risk of starvation, so to overcome this risk, the agriculture field should adapt to the IoT. Also, the IoT can pre-predicts climatic weather conditions. This will helps us to overcome all the irrigation problems and support the farmers to increase the crop yield gradually. Data mining is an area that has some rapid developments recently due to some innovations in software and hardware technology that leads to accessibility of dissimilar types of data [2]. This is true particularly for the text data case, wherever the hardware and software platforms development for the social and web networks has enabled quick formation of huge repositories of diverse data kinds [3, 4]. Specifically, the web is a technical enabler that augments the formation of huge quantity of text content through various users in the form which is simple to process and store. The rising quantities of available text data from various applications have formed a requirement for algorithmic design advances that can discover fascinating pattern from the data in a scalable and dynamic manner [5, 6]. Decision Tree-dependent classification of text does not suppose sovereignty amongst its attributes as such in Naïve Bayesian. The Decision tree perform better since a text recognition; though it become hard to make a classifier intended for large feature number [7]. At the classification time, correspondence among different features vectors class and unknown text document feature vector was intended, and the text was assigned to the class that has maximum similarity [8].

II. OVERVIEW ON WSN AND INTERNET OF THINGS

Nowadays, there is a need to implement novel ideas in the field of agriculture because of food demand conditions. The central concept of the IoT in agriculture is the improvement of the yield at a low cost [9, 10]. The IoT can act as a catalyst in the field of agriculture. There will be several types of research that can be carried out depending upon the IoT implementation in the agriculture field [11, 12]. As a result of the study, the researchers concluded that there would be a 20% increase in the growth of the agricultural industry after the implementation of the IoT [13, 14].

The agricultural industry focuses on IoT technology because of its live benefits like low input but high output within a short period [15, 16]. This IoT can make farming very smart and made it high-tech through the implementation of the novel IoT devices in the agricultural field [17, 18]. Due to the IoT utilization in the irrigation field, the manual, or else, the physical work get decreased. The sensors are implemented on the irrigation fields and the monitored data can be stored in the memory which can utilized at any time for further

process. Advanced sensors are used for monitoring. The implemented sensors are connected to the cloud network. The data obtained through the sensors are more accurate and precise. The data attained from the sensors are then stored in IoT cloud environment for further prediction process.

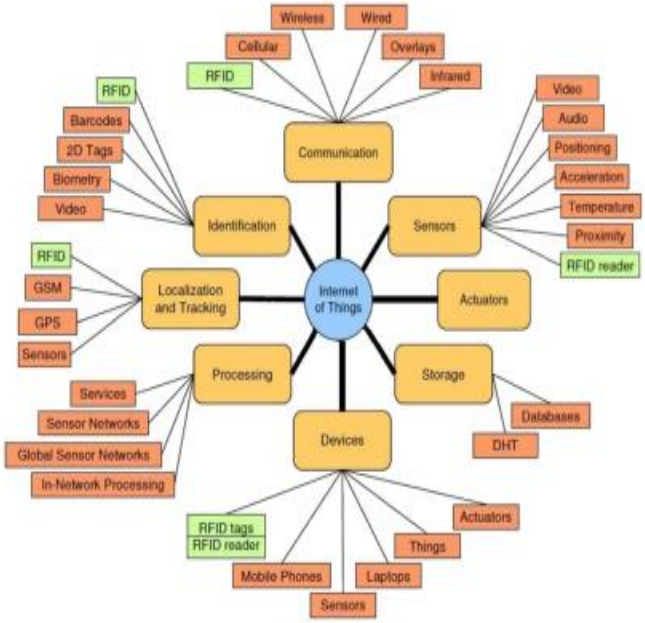

Figure 1 IoT tools

The IoT can bring out lots of benefits to the farmers like water monitoring, water usage limitation, crop monitoring, pre-prediction of the environmental impact, less manpower, increase in crop productivity, improved harvesting yield, proper field, and management [19, 20]. Then the data collection can take place. Before sowing the seeds, this kind of monitoring can help the farmers to select the crops depends upon the environment. They can collect the data about the field, and then they can choose the right crop which can uphold in a particular environment. The complete IoT setup monitors the whole field and takes a survey on the environmental factors if it can find any abnormality means an alerting condition was developed.

III. DATA AGGREGATION IN TERRESTRIAL WSN

Data aggregation in IoT for instance WSN was of very significant, as in IoT, the varied information was gathered from the varied source and further energy was required so as to transfer information/data. One such solution in this case for decreasing the energy is to aggregate and process information before sending data, and at this time, summarized and aggregated data were sent.

Wireless Sensor Networks (WSNs) incorporate of tiny devices referred as sensor nodes that having the known characteristics such as less battery power, limited storage capacity , and low computational capacity that are deployed to sensing the data collection and data communication. Conversely, the sensor nodes route the data collection through their intermediate nodes that are connected through wireless devices, to send the data to the receiver node. Maung prathub et al. (2019) [21] develop an automatic system for watering the crops in the field using the WSN. Balakrishna & Rao (2019) [22] publish a study report on using IoT agriculture farm monitoring. Here in this paper, we try to reduce the water level monitoring problems in the irrigation land. With the help of the IoT, the farmers can easily monitor the level of the water in the water tank that can make irrigation more effective. The IoT can monitor every step in the agriculture ie, from the seed sowing process to the process of the harvesting. This is a high advancement in the agricultural industry. IoT can create a massive revolution in the field of agriculture.

Mahalakshmi et al. (2020) [23] develop an automatic crop monitoring system with the use of the sensors. The sensors or else any other IoT devices get integrated for improving the agricultural practices. The sensors are implemented on the irrigation fields, and the monitored data can be stored in the memory, which can be utilized at any time for further process. Advanced sensors are used for monitoring. The implemented sensors are connected to the cloud network. The data obtained through the sensors are more accurate and precise.

Nayar et al. (2016) [24] develop a smart system that can automatically monitor the live parameters like temperature, lighting, etc. by implementing IoT technology. Wu, Q et al. (2017) [25] develop an intelligent system for monitoring the crop field. Wolters et al. [26] develop a novel method that can make farming very smart.

Gondchawar (2016) [27] develop a smart irrigation system by using the internet of things. Vidya Devi, et al. (2013) [18] put effort into developing the automatic agriculture monitoring system to make farming a modern one.

Sen,et al. (2017) [28] develop a novel smart irrigation system. This creates a peaceful revolution among the farmers. The sensors can monitor humidity, temperature, soil moisture, etc. and gave crystal clear concurrent surveillance results.

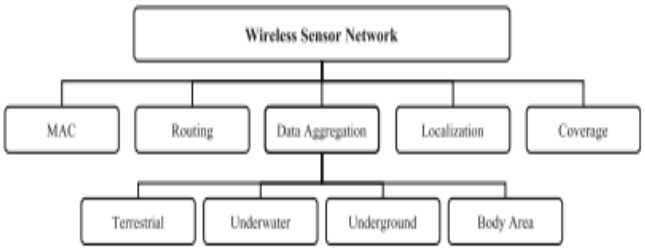

Figure 2 WSN Data aggregation

IV. BIG DATA AGGREGATION IN WIRELESS SENSOR NETWORKS

[29] stated the cluster-based data deduction used to group the data records, so the data records that exhibit closer characteristics grouped into one cluster. If those more intimate data records get deduced, the level of generalization attains optimal, and the data loss gets minimized. The k-anonymity technique depends on the clustering approach to data privacy preservation in the IoT devices to balance the exchange between data privacy and data utility.

[30] stated an efficient preventing the data privacy through the following techniques such as modification, detection, and swap used to retain the original database. The raw data evolve from the protected data for better k-means clustering to attain the privacy preservation technique. The noise added to a cluster with a lesser the *Reversible Privacy-Preserving k-means Clustering (kRPP)* of the total distance.

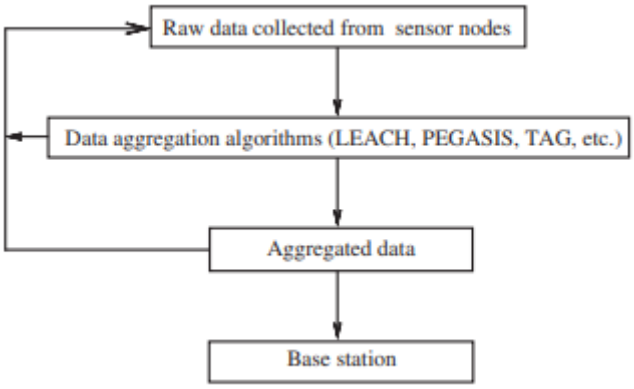

Figure 3 Data aggregation scheme architecture

[31] stated the clustering of privacy preservation methods used to protect the unnamed database against identity confession of sensitive data, confession attribute, confession linkage, and equivalent attacks. It significantly reduced the data loss—the fuzzy c-means clustering technique used to initiate the balanced clusters at least k members. The firefly algorithm used to enumerate the multi-objective function, satisfy the anonymity issues and reduced the clustering error rate.

[32] allowed the high-speed train movement intended for possessions of the network in town traffic based framework. The communication competence is developed for WSN. The planned sensor networks for an in sequence of data matching technique. To combine the observing data, the association in the middle of multi-sensors is worn in the network coverage of the sensing scheme to get better the pledge and accuracy of the scheme. In traffic flood recognition, the hypothesis is practical to vehicle type gratitude. The clear rewards over the alike LEACH protocol offer the simulation consequences demonstrate that the planned information fusion system of ordinary nodes in conditions of energy expenditure and synthesis accuracy.

[33] proposed the WSN, big data, and IoTs for ecological observing system for the weather transform on feral animals and also evolve the discovery reaction to weather transform. The surplus predators and humans from the spring place in an electric fence are worn to organize and animals progress management to limit the intruder. For ID and Wireless sensor network-based scheme observation is obtainable to reduce the person/creature dispute. Since these enormous animals journey for extremely extended distance, it is an issuing one to watch the animal observing system. In existing system, the main issue is the network limitation or there is no network coverage.

[34] presented the outbound and inbound, storeroom position and vehicle health observing. For information organization process of traditional logistics operations by distribution of vehicle healthcare monitoring management are implemented. The combination of WSN and big data analysis enabled the implementation and

monitoring system, which is tested. The test consequences establish that the scheme has convinced deployment orientation implication to improve the informationization to a certain extent and intellect of logistics organization.

[35] stated the anonymization method for the best data loss with the original database. The significant anonymization degree and retain much of the original graph properties. The limited anonymization includes anonymity with k-auto-anonymity, k-degrees, anonymity with k-neighborhood, and anonymity with 'k' as safe. The anonymization approach used to remove the amount of data and displays the multiple structural properties of data practically with the different link re-identification strategies. The anonymization method has a k-anonymity and L-diversity privacy model to enumerate efficient data transformation over the restricted constraints as conjunctive and positive. This improves the scalability, reliability, and sustained less data/information loss from the original database.

Table 1 data aggregation mechanisms Overview

Paper	Year	Node Type	Network Type	Algorithm	Aggregator
Kumar et al. [36]	2019	Homogeneous	Tree	Distributed	Fixed
Sarangi et al. [37]	2019	Homogeneous	Cluster	Distributed	Mobile
Yadav et al. [38]	2019	Homogeneous	Flat	Distributed	Fixed
SreeRanjani et al. [39]	2018	Homogeneous	Flat	Distributed	Mobile
Khriji et al. [40]	2018	Homogeneous	Cluster	Distributed	Fixed

V. DATA AGGREGATION ISSUES AND CHALLENGES

During the data aggregation process, there were some issues which should be conquered. It should be complex for conquering as shown from the related work analysis. A few significant things among them were as follows:

1. Data redundancy: it senses sensor nodes like a data type in addition to assured belongings as similar procedures, therefore the sink node collect information which is unnecessary. Therefore, it is energy, other resources and time wastage.
2. Delay: In assured cases information from beyond nodes appear behind at the root node or sink accordingly of which the process of aggregation start not on time. Additional, at the intermediate aggregations level increase the delay time further.
3. Accuracy: There were two accuracy issues types. Initially, the function of aggregator is an estimate task; hence, loss of accuracy on the data forwarding state. Next, there should be a node that is compromised for sending inappropriate / false data to the node of aggregator. The node of Aggregator does not guarantee its accuracy and process it.
4. Traffic load: In the scenarios assured, the node of aggregator is overburden. It happens but the balancing of load is not made or the clusters are of uneven sizes.
5. Aggregation freshness: Information from comparable frames must be old stored data, aggregated or data aggregation from several frames as of diverse time must be avoided since it loses originality.

VI. CONCLUSION

In this manuscript, various approaches have been surveyed for the aggregation and reduction of data with the use of aggregation. The techniques that were surveyed were based on few critical impacts like type of network, heterogeneity of the node, aggregator node mobility, and type of algorithm. It was recognized that most techniques attempt to enhance the latency or efficiency of the system. Various existing solutions are there for the data aggregation in WSN. But every existing method should possess an advantage and disadvantage. From this survey, it was concluded that still the data aggregation technique should have some open issues apart from the various solutions presented. Thus, there is a need to implement some effective techniques.

REFERENCES

[1] S. A. Dehkordi, K. Farajzadeh, J. Rezazadeh, R. Farahbakhsh, K. Sandrasegaran, and M. A. Dehkordi, "A survey on data aggregation techniques in IoT sensor networks," *Wireless Networks*, vol. 26, pp. 1243-1263, 2020.

[2] O. C. Abikoye, S. O. Omokanye, and T. O. Aro, "Text Classification Using Data Mining Techniques: A Review," *Computing & Information Systems*, vol. 22, 2018.

[3] M. Kang, J. Ahn, and K. Lee, "Opinion mining using ensemble text hidden Markov models for text classification," *Expert Systems with Applications*, vol. 94, pp. 218-227, 2018.

[4] K. Kowsari, D. E. Brown, M. Heidarysafa, K. J. Meimandi, M. S. Gerber, and L. E. Barnes, "Hdltex: Hierarchical deep learning for text classification," in *2017 16th IEEE international conference on machine learning and applications (ICMLA)*, 2017, pp. 364-371.

[5] L. Jiang, C. Li, S. Wang, and L. Zhang, "Deep feature weighting for naive Bayes and its application to text classification," *Engineering Applications of Artificial Intelligence*, vol. 52, pp. 26-39, 2016.

[6] H. Rahman, N. Ahmed, and I. Hussain, "Comparison of data aggregation techniques in Internet of Things (IoT)," in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 2016, pp. 1296-1300.

[7] T. Pranckevičius and V. Marcinkevičius, "Comparison of naive bayes, random forest, decision tree, support vector machines, and logistic regression classifiers for text reviews classification," *Baltic Journal of Modern Computing*, vol. 5, p. 221, 2017.

[8] S. Qaiser and R. Ali, "Text mining: use of TF-IDF to examine the relevance of words to documents," *International Journal of Computer Applications*, vol. 181, pp. 25-29, 2018.

[9] W. Ding, X. Jing, Z. Yan, and L. T. Yang, "A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion," *Information Fusion*, vol. 51, pp. 129-144, 2019.

[10] J. Guo, R. Chen, and J. J. Tsai, "A survey of trust computation models for service management in internet of things systems," *Computer Communications*, vol. 97, pp. 1-14, 2017.

[11] P. Visu and V. Murugananthan, "A dependable smart security system for dominant attacks on agriculture crop fields," *International Journal of Engineering & Technology*, vol. 7, pp. 195-197, 2018.

[12] B. Pourghebleh and N. J. Navimipour, "Data aggregation mechanisms in the Internet of things: A systematic review of the literature and recommendations for future research," *Journal of Network and Computer Applications*, vol. 97, pp. 23-34, 2017.

[13] I. Ali, E. Khan, and S. Sabir, "Privacy-preserving data aggregation in resource-constrained sensor nodes in Internet of Things: A review," *Future Computing and Informatics Journal*, vol. 3, pp. 41-50, 2018.

[14] F. Shi, Q. Li, T. Zhu, and H. Ning, "A survey of data semantization in internet of things," *Sensors*, vol. 18, p. 313, 2018.

[15] S. Satarkar, S. Agrawal, S. Makwana, and P. Joshi, "Decision Support System for Smart Agriculture using IoT."

[16] M. Gayatri, J. Jayasakthi, and G. A. Mala, "Providing Smart Agricultural solutions to farmers for better yielding using IoT," in *2015 IEEE Technological Innovation in ICT for Agriculture and Rural Development (TIAR)*, 2015, pp. 40-43.

[17] N. K. Singh, D. S. Tomar, and A. K. Sangaiah, "Sentiment analysis: a review and comparative analysis over social media," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-21, 2018.

[18] G. M. Kumari and D. V. Devi, "Real Time Automation and Monitoring System for Modernized Agriculture," *International Journal of Review & Research in Applied Sciences and Engineering vol3*, 2013.

[19] S. Rukhmode, G. Vyavhare, S. Banot, A. Narad, and R. Tugnayat, "IOT based agriculture monitoring system using wemos," *no. March*, pp. 14-19, 2017.

[20] M. M. BASAM, S. S. KUMAR, and S. S. SELVAN, "IOT BASED MONITORING SYSTEM IN SMART AGRICULTURE."

[21] J. Muangprathub, N. Boonnam, S. Kajornkasirat, N. Lekbangpong, A. Wanichsombat, and P. Nillaor, "IoT and agriculture data analysis for smart farm," *Computers and electronics in agriculture*, vol. 156, pp. 467-474, 2019.

[22] G. Balakrishna and M. N. Rao, "Study Report on Using IoT Agriculture Farm Monitoring," in *Innovations in Computer Science and Engineering*, ed: Springer, 2019, pp. 483-491.

[23] J. Mahalakshmi, K. Kuppusamy, C. Kaleeswari, and P. Maheswari, "IoT Sensor-Based Smart Agricultural System," in *Emerging Technologies for Agriculture and Environment*, ed: Springer, 2020, pp. 39-52.

[24] A. Nayyar and E. V. Puri, "Smart Farming: IoT Based Smart Sensors Agriculture Stick for Live Temprature and Moisture Monitoring using Arduino Cloud Computing & Solar Technology," in *Conference: The International Conference on Communication and Computing Systems (ICCCS-2016)*, 2016.

[25] Q. Wu, Y. Liang, Y. Li, and Y. Liang, "Research on intelligent acquisition of smart agricultural big data," in *2017 25th International Conference on Geoinformatics*, 2017, pp. 1-7.

[26] N. Katyal and B. J. Pandian, "A Comparative Study of Conventional and Smart Farming," in *Emerging Technologies for Agriculture and Environment*, ed: Springer, 2020, pp. 1-8.

[27] N. Gondchawar and R. Kawitkar, "IoT based smart agriculture," *International Journal of advanced research in Computer and Communication Engineering*, vol. 5, pp. 838-842, 2016.

[28] S. Sen and B. Madhu, "Smart agriculture: a bliss to farmers," *Int. J. Eng. Sci. Res. Technol*, 2017.

[29] K. Arava and S. Lingamgunta, "Adaptive K-anonymity approach for privacy preserving in cloud," *Arabian Journal for Science and Engineering*, pp. 1-8, 2019.

[30] C.-Y. Lin, "A reversible privacy-preserving clustering technique based on k-means algorithm," *Applied Soft Computing*, vol. 87, p. 105995, 2020.

[31] R. K. Langari, S. Sardar, S. A. A. Mousavi, and R. Radfar, "Combined fuzzy clustering and firefly algorithm for privacy preserving in social networks," *Expert Systems with Applications*, vol. 141, p. 112968, 2020.

[32] H. Kong and B. Yu, "An improved method of WSN coverage based on enhanced PSO algorithm," in *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, 2019, pp. 1294-1297.

[33] A. Yahya, "Climate: Environmental Monitoring Using Wireless Sensor Network System," in *Emerging Technologies in Agriculture, Livestock, and Climate*, ed: Springer, 2020, pp. 67-98.

[34] J. Jiang, X. Zhu, G. Han, M. Guizani, and L. Shu, "A Dynamic Trust Evaluation and Update Mechanism based on C4. 5 Decision Tree in Underwater Wireless Sensor Networks," *IEEE Transactions on Vehicular Technology*, 2020.

[35] M. Kiranmayi and N. Maheswari, "A Review on Privacy Preservation of Social Networks Using Graphs," *Journal of Applied Security Research*, pp. 1-34, 2020.

[36] S. Kumar and H. Kim, "Energy efficient scheduling in wireless sensor networks for periodic data gathering," *IEEE access*, vol. 7, pp. 11410-11426, 2019.

[37] K. Sarangi and I. Bhattacharya, "A study on data aggregation techniques in wireless sensor network in static and dynamic scenarios," *Innovations in systems and software engineering*, vol. 15, pp. 3-16, 2019.

[38] S. Yadav and R. S. Yadav, "Redundancy elimination during data aggregation in wireless sensor networks for IoT systems," in *Recent Trends in Communication, Computing, and Electronics*, ed: Springer, 2019, pp. 195-205.

[39] N. SreeRanjani, A. Ananth, and L. S. Reddy, "An energy efficient data gathering scheme in wireless sensor networks using adaptive optimization algorithm," *Journal of Computational and Theoretical Nanoscience*, vol. 15, pp. 3456-3461, 2018.

[40] S. Khriji, G. V. Raventos, I. Kammoun, and O. Kanoun, "Redundancy elimination for data aggregation in wireless sensor networks," in *2018 15th International Multi-Conference on Systems, Signals & Devices (SSD)*, 2018, pp. 28-33.