FAULT-TOLERANT AND EFFICIENT DATA AGGREGATION TECHNIQUES IN RESOURCE-CONSTRAINED WSN ENVIRONMENTS

Vijaysinh Uttamrao Bansude
Research Scholar
Dept Of ECE
SJJTU Rajasthan

Dr. Rahulkumar Budania
Supervisor
Dept Of ECE
SJJTU Rajasthan

Dr. S. T. Shirkande
Co-Supervisor
Dept Of ECE
SJJTU Rajasthan

ABSTRACT

Data aggregation is a fundamental process that enhances data analysis and decision-making across various industries. Organizations can gain valuable insights, improve efficiency, and make informed decisions by collecting and summarizing data from multiple sources. Whether it's used in business intelligence, healthcare, marketing, finance, or environmental studies, the benefits of data aggregation are far-reaching. Our experimental results show that our approach reduces the communication cost, transmission cost, energy consumption at the nodes and cluster heads, and delivers highly-refined and fused data to the base station. Embracing this process is essential for any organization looking to harness the power of data. However, the data privacy and security are challenging the proliferation of clustering wireless sensor network (CWSN) due to its highly constrained resources and violably deployed environments, which make it infeasible to directly apply traditional cryptography and therefore vulnerable to various attacks. This study proposes a scheme that provides efficient privacy-preserving data fusion as well as malicious data tolerance by mining concealed data within groups. And the dynamically organized groups in each cluster improves resilience against large number of node compromise comparing with the existing data aggregation schemes.

Keywords: *Data aggregation, clustering wireless sensor network (CWSN), decision-making, communication cost, environments.*

INTRODUCTION

Data aggregation is becoming more and more important as IoT applications become more common and diverse, and as WSN becomes more significant in domains like industrial automation, healthcare, and environmental monitoring. This process is not just a minor detail; it is the crucial element that supports the long-term viability and improvement of these complex and interconnected ecosystems. The meticulous organization of data through aggregation guarantees that these networks maintain efficiency, reliability, and adaptability to the changing requirements of the digital era. Currently, in a time characterized by remarkable digital change and interconnectivity, WSN and the IoT are leading the way in technological advancement. Several sensor nodes make up these networks, which collaborate to gather and process environmental data. Data aggregation plays a vital role in this complex environment by coordinating the gathering, processing, and transfer of sensor data. The major goal of data aggregation in WSNs and IoT is to find a careful balance between conserving important energy resources, minimising the cost of data transmission, and making sure the data collected is of good quality. With the ongoing evolution and expansion of networks, the importance of data aggregation is increasingly evident. The constantly changing and limited resources of WSN and IoT environments have sparked the quest for inventive protocols, techniques, and strategies to tackle the complex challenges that emerge. As a result, a significant amount of literature has emerged, providing various viewpoints and solutions to the complex challenges presented by data aggregation in these networks.

LITERATURE REVIEW

Babedi Betty Letswamotse (2018) Short-range communication networks are increasingly using wireless sensor networks, or WSNs. Over the last several decades, their application acceptability to mission-critical systems has increased dramatically because to their inexpensive installation costs and simplicity of deployment. Beyond their computational capabilities, these technologies have many drawbacks. Therefore, several previous research methodologies have attempted to enhance their processing, communication, robustness, and other capacities. One strategy is to build WSNs that provide a certain degree of QoS guarantee in their applications, as the dynamic architecture of WSNs makes it challenging to deliver the proper Quality of Service (QoS) outcomes. This is challenging, however, since effective resource management is essential to achieving the required quality of service in WSNs. A review of QoS provision in WSNs, specifically for mission-critical applications, is presented in this work. The implementation of QoS metrics, QoS difficulties, and enhanced communication possibilities in WSNs are the main topics of the review. In order to increase network performance, the article also suggests Software Defined Wireless Sensor Networking (SDWSN) techniques for effective resource management and assured QoS support.

Valério Rosset (2017) One of the biggest problems with (WSNs) is finding a way to balance cost-effective energy conservation with reliable data transfer. This study creates a new bio-inspired routing protocol called CB-RACO using the Ant Colony Optimisation (ACO) meta-heuristic and the Label Propagation (LP) approach for finding communities that is cheap and easy to use. By using swarm intelligence to route data across communities, CB-RACO establishes communities in WSNs and balances energy usage. As a result, CB-RACO needs little overhead and memory to construct and maintain routing paths. To further guarantee high data delivery reliability, CB-RACO employs a data retransmission mechanism based on community acknowledgements. We ran large-scale simulations of CB-RACO based on parameters like energy consumption, delivery delay, and excellent put. The findings suggest that the proposed method may be much better than ant-based solutions that don't depend on community design.

Rajashekhara (2017) An essential component of WSN is data aggregation. since it makes sure that the user gets the environmental data that the base station collects. Compared to the more often discussed routing and energy challenges, there are other underlying concepts that need greater attention in order to assure effective data aggregation. We looked at existing data aggregation techniques, focussing on data correlation methods in particular, and found that a lot more work is required in this area. We find that there is a very small number of research publications about existing correlational-based data aggregation methods. It was also investigated that these methods still pay little attention to issues such improper benchmarking, computational complexity, and data quality. This study describes all of the open problems that need to be investigated with a specific emphasis on improving the data quality and reliability in the wireless sensor network aggregation process.

Amandeep Kaur (2016) Because of its numerous uses in industries including environmental monitoring, military surveillance, healthcare, and disaster assistance, or WSNs, have attracted a lot of attention. Energy is the main constraint for WSN since the sensor nodes have a finite and irreplaceable power supply. The most often used topology control technique for increasing WSN scalability and lowering energy usage is clustering. In this article, we introduced DFCA, a distributed fault-resistant clustering technique that creates clusters according to the cluster heads' cost function. Furthermore, because of the unexpected failure of the cluster head, we provide a distributed run-time recovery of the sensor nodes from the troubled cluster. The experimental findings show how effective the suggested approach is. This problem may be solved by either using a backup cluster or by prefixing the root information.

Geetika Dhand (2016) The energy required to process raw data varies depending on the data aggregation strategy. The requirements of the application and the proportionate energy savings from this approach determine which data aggregation technique is best. Since several sensor nodes often detect the same events, there will be a significant amount of data duplication. Many applications, in the meanwhile, use more sensors than are required to accurately identify the intended phenomena. In addition to explaining the importance of data collecting, this paper examines a number of hierarchical clustering techniques.

Data Aggregation: Make Your Analysis More Efficient

The majority of team members in a contemporary organisation will use aggregated data to solve issues and reach conclusions. Spreadsheets continue to be the preferred tool for data aggregation for many corporate users. From the spreadsheet interface, they may quickly compute and view sums, averages, and other data summaries.

However, a spreadsheet is insufficient for data aggregation at the level of big data, data science, and sophisticated statistical analysis. In this situation, data experts need to optimise backend query performance against billions of entries and streamline analysis using their coding expertise and contemporary data stack.

The primary advantages of data aggregation are the same regardless of the application setting. Data aggregation increases the efficacy of data analysis and aids all users in understanding the data they are working with.

To link real-time data from all of your systems to Sheets and expedite your data analysis, try Coefficient's free plan right now.

Introduction to wireless sensor networks (WSNs) a novel kind of wireless networking technique is gaining popularity in commercial and military settings. A network of autonomous sensor devices dispersed across a wide region is known as a or WSN for short. They are used to measure changes in the environment or in physical factors. A(WSN) is composed of several small sensor nodes that communicate and exchange information. These nodes gather and transmit environmental data to a base station, including temperatures, pressures, humidity, and contaminants. The latter either communicates the data to a network that is connected to it or initiates an action or alert, depending on the kind and volume of data being monitored.

Common uses include tracking the movement of people and animals in woods and along borders, monitoring the weather and forests, monitoring the battlefield, and physically measuring environmental parameters including pressure, temperature, vibration, and pollution. Like wireless local area networks (WLANs), they employ the air as their wireless transmission channel. IEEE 802.11 and other standard access protocols are available to ensure that nodes in a local area network are properly communicating with one another. Nevertheless, WSNs cannot directly use this or any of the other protocols. The main difference is that sensors have a small power source, like a battery, that runs out rapidly. This is different from devices that are part of local area networks. So, it is important to design new MAC protocols that are conscious of energy. There is clearly a difference between a WSN and a WLAN since a WSN has less resources.

Components of a wireless sensor node

The sensor node is a crucial component of a WSN. This little device has the ability to map or save data about its immediate surroundings. The cost of these devices is constantly falling as a result of advancements in semiconductor technology. The following are the primary parts of these little gadgets.

microcontroller. Despite being so small, this computer-on-a-chip is able to perform a variety of complex duties, such as managing the operations of other linked devices. A CPU, a RAM memory, and related peripherals make up a microcontroller in general. Many devices available today are capable of doing the same tasks as microprocessors. Among them are field-programmable gate arrays (FPGAs), application-specific integrated circuits (ASICs), and digital signal processors (DSPs). Microprocessors are the ideal option for tiny to very compact embedded systems because of their moderate to high processing capabilities and low power consumption, even if each of these devices has pros and downsides.

Classification of sensor networks

Numerous applications with diverse needs have emerged as a result of wireless sensor networks' quick growth. In order to handle these changing needs, a number of distinct network architects have created protocols for various network levels. Here, we highlight some of the key distinctions across sensor networks, despite the fact that there are several methods to classify their architecture.

Data sinks: The characteristics of the data sink or sinks are among the most important components of sensor networks. In certain situations, the end user or users may be mobile access points that periodically collect data, or they may be permanently anchored within the sensor network. This distinction could be significant since, in the second scenario, effective distributed data storage techniques might work better.

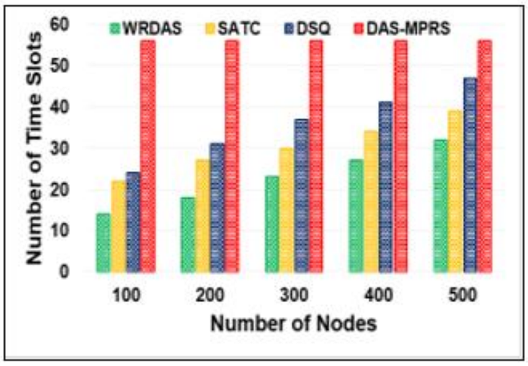
Mobility of sensors: The type of sensor being organised may serve as the basis for a different classification of sensor networks. Although sensors are typically thought of as fixed, several contemporary sensor networks initiatives, such as Mobile sensor nodes are included into ZebraNet. In military applications, sensors may be affixed to unmanned aerial vehicles (UAVs) or troops' bodies or uniforms to connect to a structured sensor network.

RESEARCH METHODOLOGY

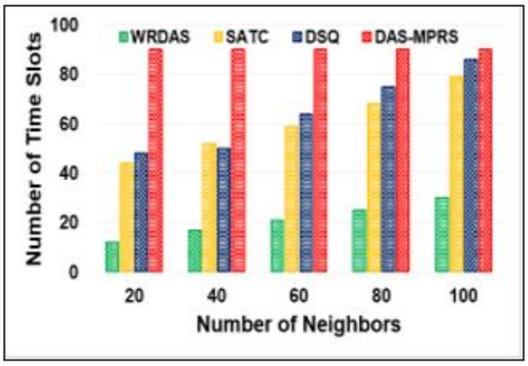
The subsequent hop's Q-value is adjusted after reward computation. The node to be investigated is the subsequent hop. Routes are constructed from all sources to the sink in this manner. After discovering all the routes, this aggregation tree construction component checks the routes and identifies the common aggregation points and perform data aggregation at these intermediate nodes. The overlapping paths resulted in the creation of a reverse multicast data aggregation tree, where the sink served as the root. This component generates a conflict-free transmission schedule for the nodes of the aggregate tree built by the aggregation tree construction component. The weight of each leaf is determined by its distance from the sink, number of neighbors, and remaining node energy. Rankings are provided after the weight calculation, and the nodes are organised in a non-increasing weight order. The node with the biggest id is given the highest rank among nodes with the same weight. The leaf nodes that satisfy the conflict-free criteria are given the same time slot. The next time slot is used for nodes that are not scheduled in the current one. It seems sense to eliminate the time slot-assigned leaf nodes from the tree. The aforementioned procedure is repeated for the remaining tree structures until all nodes have time slots assigned to them. The data that the source nodes see is combined with the secrets to form the plaintext data. The integrity information that is encoded in the data produced by the source nodes is known as the secret.

RESULTS AND DISCUSSIONS

The number of scheduling windows needed to allot time slots to every node in the aggregate tree is assessed by this measure. Graph 1 shows the maximum amount of time slots that each scheduling system may use. With the exception of DAS-MPRS, which deviates from this pattern, it is clear from Graph 1 (a) that in scenario 1, as the number of nodes increases, so does the amount of time slots used.



(a) Scenario 1



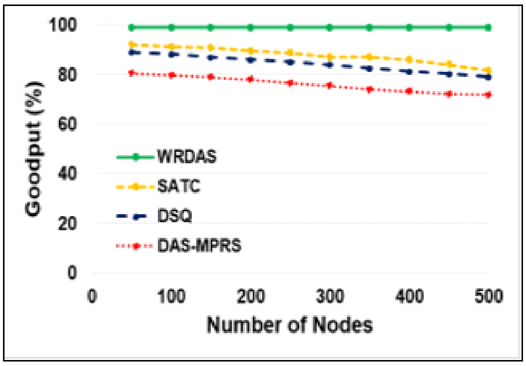
(b) Scenario 2

Graph 1: Time Slot Allocation

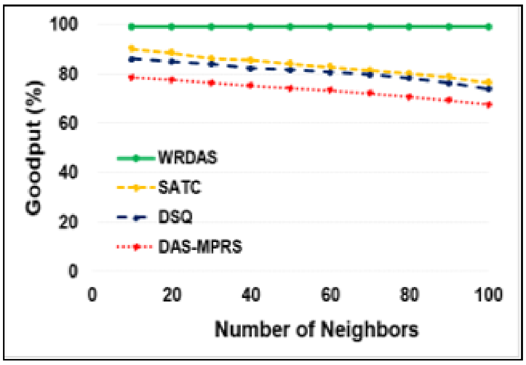
Because the number of colours used to colour the cells is fixed in DAS-MPRS, the number of time slots stays constant. This fixed coloring scheme limits the flexibility of scheduling, leading to a uniform time slot count regardless of the number of nodes. Additionally, it occupies more spaces than DSQ, SATC, and WRDAS. Compared to SATC and DSQ, WRDAS uses fewer maximum time slots because slot assignment is based on computed ranks, allowing efficient reuse of the same time slot among more nodes. SATC, which considers both the forward degree and collision degree during time slot allocation, requires fewer slots than DSQ.

The quantity of valuable data bits that made it to the sink node is known as goodput.

Also known as network throughput, Graph 2 displays the outcomes. The outcomes of scenario 1 demonstrate that the number of nodes in WRDAS has no effect on goodput. The efficient scheduling of nodes based on their residual energy, distances from the sink, along with the number of neighbours is what causes this stability, which results in the highest goodput performance. When compared to other algorithms, DAS-MPRS performs poorly, whereas SATC and DSQ produce roughly same goodput.



(a). Scenario 1

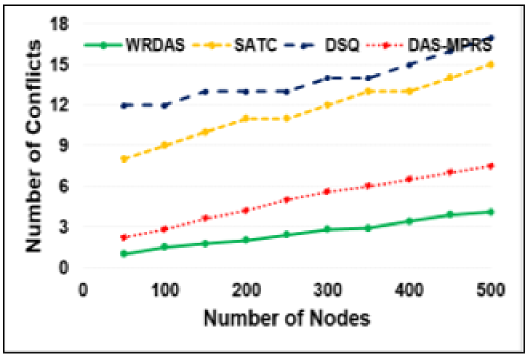


(b). Scenario 2

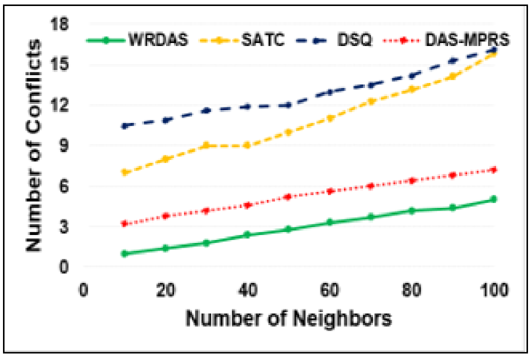
Graph 2: Goodput

Given that WRDAS schedules nodes using the weighted average approach, Graph 2 (b) shows that its performance is comparable to scenario 1, but other methods' goodput is lower than scenario 1. The explanation is that as the number of neighbours rises, node conflicts grow as well, leading to lower output.

It speaks about the quantity of nodes that the transmission from the present node may affect. Conflict is one of the metrics which decides the overall performance of WSNs. Conflicts occur when the network nodes are not well scheduled. In data aggregation, conflicts play a vital role as they impact the aggregation results.



(a). Scenario 1



(b). Scenario 2

Graph 3: Conflicts

Graph 3 shows the statistics showing the amount of conflicts relative to the number network neighbours and nodes. Graph 3 (a) shows that the number of conflicts rises in proportion to the total amount of nodes. As WRDAS schedules the nodes using two hop neighborhood, it greatly reduces the conflicts among the nodes. There are less conflicts in WRDAS even when the number of nodes rises.

CONCLUSIONS

This method improves the efficiency of the underlying collection of information process while addressing some of the drawbacks of the conventional K-means. Nevertheless, data aggregation's effectiveness and performance are subpar for the suggested approach. A straightforward function-based data aggregation technique aimed at cluster-based WSNs was presented. Their approach uses the Min and Max functions to eliminate data redundancy at the node level. After comparing the stored Min value with the incoming sensor data, if the new value is smaller, it replaces the old Min in the node's buffer. The same iterative process applies for updating the Max value. This method effectively reduces redundancy in the collected data by maintaining only the essential minimum and maximum readings. eliminated, and a portion of the locally aggregated data is sent to the cluster leaders. NATURE, a meticulously crafted trust-based method for identifying misconduct in clustered Industrial Wireless Sensor Networks (IWSNs), is presented in this study. Data confidence and communication trust are the two main trust metrics used by NATURE. It evaluates communication trust by tracking both primary interactions & aggregate data from neighboring nodes.

REFERENCES

1. *Babedi Betty Letswamotse (2018), "Software Defined Wireless Sensor Networks (SDWSN): A Review on Efficient Resources, Applications and Technologies", Journal of Internet Technology, ISSN:2079-4029, Volume. 19, issue.5*
2. *Valério Rosset (2017), "Enhancing the Reliability on Data Delivery and Energy Efficiency by Combining Swarm Intelligence and Community Detection in Large-scale WSNs", Expert Systems with Applications, ISSN: 1873-6793, Vol. 78, issue. (4), DOI: 10.1016/j.eswa.2017.02.008*
3. *Rajashekhara (2017), "Exploring an Effectiveness & Pitfalls of Correlational-based Data Aggregation Approaches in Sensor Network", International Journal of Wireless and Microwave Technologies, ISSN :2076-9539, vol.7, issue.2, https://doi.org/10. 5815/ijwmt.2017.02.05*
4. *Amandeep Kaur (2016), "Enhanced Routing Scheme Using Dynamic Clustering for Energy Reduction in WSN", International Journal of Advanced Research in Computer Science, vol.7, issue.6*
5. *Geetika Dhand (2016), "Data Aggregation Techniques in WSN: Survey", Procedia Computer Science, ISSN:1877-0509, Volume.92, Pages.378-384, https://doi.org/10.1 016/j.procs.2016.07.393*
6. *Ameya S. Bhatlavande (2015), "Energy Efficient Approach for In-Network Aggregation in Wireless Sensor Networks", International Journal of Current Engineering and Technology, ISSN:2277–4106, Vol.5, issue.4,*
7. *C. Sahaya Kingsly, M. E (2013), "Critical Study on Constraints in Wireless Sensor Network Applications", International Journal of Engineering Research & Technology, ISSN:2278-0181, Vol.2, Issue.7,*
8. *Chamran Asgari (2012), "A New Approach to the Data Aggregation in Wireless Sensor Networks", IJCSI International Journal of Computer Science Issues, ISSN: 1694-0814, Vol.9, Issue.1*
9. *Chaudhari H.C (2011), "Wireless Sensor Networks: Security, Attacks and Challenges", International Journal of Networking, ISSN: 2249-2798, Volume. 1, Issue.1, pp-04-16.*
10. *Claude Castelluccia (2009), "Efficient and Provably Secure Aggregation of Encrypted Data in Wireless Sensor Networks", ACM Transactions on Sensor Networks, ISSN: 1550-4867, vol.5, issue. (3), DOI:10.1145/1525856.1525858*