ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 3 | June 2024

www.ijirid.in

A Hybrid Spatio-Temporal and Machine Learning-Based Data Aggregation Model for Redundancy Reduction in Wireless Sensor Networks

Sagar Padiya[1], Aniket Thakur[2]

[1]Assistant Professor, IT, Shri Sant Gajanan Maharaj College of Engineering, Shegaon, India

[2]Assistant Professor, IT, Sipna College of Engineering & Technology, Amravati, India

Abstract: *Wireless Sensor Networks (WSNs) suffer from excessive data redundancy due to dense deployment and overlapping sensing regions. Traditional data aggregation techniques reduce redundancy but often fail to balance energy efficiency, accuracy, and adaptability. This paper proposes a Hybrid Spatio-Temporal Machine Learning-based Data Aggregation (HST-MLDA) model that integrates clustering, correlation analysis, and lightweight machine learning techniques to eliminate redundant data effectively. The proposed model combines temporal filtering at the node level and spatial correlation at the cluster level, followed by classification-based redundancy elimination. Simulation-based analysis demonstrates improved network lifetime, reduced energy consumption, and enhanced data accuracy compared to traditional aggregation techniques.*

Keywords: *Wireless Sensor Networks, Data Aggregation, Data Redundancy, Clustering, Machine Learning, Energy Efficiency, etc.*

I. INTRODUCTION

Wireless Sensor Networks (WSNs) consist of spatially distributed sensor nodes that monitor environmental conditions such as temperature, humidity, and pressure. Due to dense deployment and overlapping sensing ranges, multiple nodes often capture similar data, resulting in redundancy. This redundancy leads to unnecessary data transmission, increased energy consumption, and reduced network lifetime.

Data aggregation techniques aim to reduce redundant transmissions by combining data from multiple nodes. However, conventional approaches often rely solely on spatial or temporal filtering and fail to adapt to dynamic environments. Moreover, they may compromise data accuracy while reducing redundancy. To address these challenges, this paper proposes a Hybrid Spatio-Temporal Machine Learning-based Data Aggregation (HST-MLDA) model that integrates:

- Temporal redundancy reduction at the sensor node level
- Spatial correlation at the cluster level
- Machine learning-based redundancy classification

The proposed approach aims to optimize energy efficiency without compromising data fidelity..

II. RELATED WORK

Data aggregation in WSNs has been widely studied, with approaches broadly categorized into:

2.1 Spatial Aggregation

Cluster-based protocols such as LEACH (Low-Energy Adaptive Clustering Hierarchy) reduce energy consumption by aggregating data at cluster heads. However, they often overlook temporal redundancy.

2.2 Temporal Aggregation

Temporal techniques reduce redundancy by eliminating repeated data over time. While effective, these methods fail when spatial correlation dominates.

2.3 Machine Learning-Based Approaches

Recent studies use machine learning algorithms such as decision trees and neural networks to detect redundant data. These approaches improve accuracy but may introduce computational overhead unsuitable for resource-constrained nodes.

2.4 Limitations

Existing methods typically:

- Focus on either spatial or temporal redundancy
- Lack adaptability to dynamic environments
- Do not effectively balance energy efficiency and data accuracy:

III. PROPOSED SYSTEM

3.1 System Architecture

The proposed model consists of three layers:

1. Sensor Node Layer
2. Cluster Head Layer
3. Base Station Layer

Each layer performs specific redundancy reduction tasks.

3.2 Temporal Redundancy Reduction (Node Level)

Each sensor node performs lightweight temporal filtering:

- Data is compared with previously transmitted values
- If the variation is below a threshold (ΔT), transmission is suppressed

Condition:

$$|D\text{current} - \text{Dprevious}| < \Delta T$$

This reduces unnecessary transmissions at the source.

3.3 Spatial Correlation (Cluster Level)

Cluster Heads (CHs) perform spatial redundancy elimination:

- Data from neighboring nodes is analyzed

 DOI: 10.5281/zenodo.19143838

ISSN (Online): 2583-648X
IJIRID
International Journal of Ingenious Research, Invention and Development
An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal
Volume 3 | Issue 3 | June 2024
www.ijirid.in

- Correlation coefficient (ρ) is computed

Highly correlated data (ρ > threshold) is aggregated into a single representative value.

3.4 Machine Learning-Based Redundancy Classification

A lightweight classifier (e.g., Decision Tree or Naïve Bayes) is used at the CH level:

- Input features:
  - Temporal similarity
  - Spatial correlation
  - Node distance
  - Historical variance
- Output:
  - Redundant / Non-redundant classification

This step ensures intelligent redundancy removal without compromising accuracy.

3.5 Algorithm Workflow

1. Sensor nodes collect data
2. Apply temporal filtering
3. Transmit filtered data to cluster heads
4. Cluster heads perform spatial correlation
5. Apply ML classifier for redundancy detection
6. Aggregate and forward data to base station

IV. PERFORMANCE EVALUATION

4.1 Simulation Setup

- Simulator: NS-2 / MATLAB
- Number of nodes: 100–500
- Deployment area: 100m × 100m
- Initial energy: 2 Joules per node
- Metrics:
  - Energy consumption
  - Network lifetime
  - Packet delivery ratio
  - Data accuracy

4.2 Result and Discussion

4.2.1 Energy Consumption: The HST-MLDA model reduces energy consumption by minimizing redundant transmissions at multiple levels.

4.2.2 Network Lifetime: The proposed approach extends network lifetime by 25–35% compared to traditional LEACH-based aggregation.

 DOI: 10.5281/zenodo.19143838

ISSN (Online): 2583-648X
IJIRID
International Journal of Ingenious Research, Invention and Development
An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal
www.ijirid.in
Volume 3 | Issue 3 | June 2024

4.2.3 Data Accuracy: Unlike aggressive aggregation techniques, the ML-based classification maintains high data fidelity.

4.2.4 Redundancy Reduction: The hybrid approach achieves significant redundancy elimination due to combined spatial and temporal filtering..

V. ADVANTAGES OF THE PROPOSED MODEL

- Multi-level redundancy reduction
- Improved energy efficiency
- Adaptive to dynamic environments
- Maintains high data accuracy
- Scalable for large WSN deployments

VI. LIMITED AND FUTURE WORK

Limitations

- Slight computational overhead at cluster heads
- Requires initial training of ML models

Future Work

- Integration of deep learning models for better prediction
- Real-world deployment validation
- Optimization for IoT and edge computing environments.

VI. CONCLUSION

This paper presented a Hybrid Spatio-Temporal Machine Learning-based Data Aggregation (HST-MLDA) model for reducing redundancy in Wireless Sensor Networks. By combining temporal filtering, spatial correlation, and machine learning classification, the proposed approach effectively minimizes redundant data transmission while preserving data accuracy. Simulation results demonstrate improved energy efficiency, extended network lifetime, and enhanced performance compared to traditional aggregation techniques. The model offers a promising solution for next-generation intelligent WSN systems.

REFERENCES

[1] Heinzelman, W. et al., "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," IEEE, 2000.
[2] Yick, J., Mukherjee, B., Ghosal, D., "Wireless Sensor Network Survey," Computer Networks, 2008.
[3] Krishnamachari, B., "Networking Wireless Sensors," Cambridge University Press, 2005.
[4] Akyildiz, I. F., et al., "Wireless Sensor Networks: A Survey," IEEE Communications Magazine, 2002.
[5] Rajagopalan, R., Varshney, P., "Data Aggregation Techniques in Sensor Networks," IEEE, 2006

 DOI: 10.5281/zenodo.19143838

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

www.ijirid.in

Volume 3 | Issue 3 | June 2024

[6] S. D. Padiya and V. S. Gulhane, “Analysis of Data Aggregation Methods to Avoid Data Redundancy in Wireless Sensor Networks,” in 2020 IEEE 12th International Conference on Computational Intelligence and Communication Networks (CICN), Bhimtal, India, 2020, pp. 123-129. IEEE Xplore.

[7] S. D. Padiya and V. S. Gulhane, "Eddystone-TLM frame with a lightweight frame for sensor data broadcasting," International Journal of Internet Protocol Technology, vol. 17, no. 3, pp. 133–146, May 13, 2025, doi: 10.1504/IJIPT.2024.146222.

[8] S. D. Padiya and V. S. Gulhane, “Analysis of Bluetooth Versions (4.0, 4.2, 5, 5.1, and 5.2) for IoT Applications,” in Implementing Data Analytics and Architectures for Next Generation Wireless Communications, C. Bhatt, N. Kumar, A. K. Bashir, and M. Alazab, Eds. Hershey, PA, USA: IGI Global, 2022, pp. 153-178.

