A Comparative Analysis of Elaborated Data Aggregation Methods in Wireless Sensor Networks and IoT-Based Smart Sensor Systems

Molagajjela Devaraju[1]

[1]*Assistant Professor, Dept. of ECE, GIET, Hyderabad; Ph.D. Scholar, BESTIU, A.P.*

Abstract

Wireless Sensor Networks (WSNs) are vital IoT components that enable data collection across the military, agricultural, and healthcare domains. Sensor nodes operate on constrained battery power, making energy efficiency and network longevity critical concerns. This study presents a comparative analysis of four advanced data aggregation methods for WSNs: (1) TDAC — a hybrid Tree-based Data Aggregation algorithm combining Ant Colony Optimization (ACO) and Cuckoo Search (CS); (2) an energy-aware aggregation scheme for heterogeneous initial-energy nodes; (3) Quad Tree-based clustering framework with edge-assisted UAV relay; and (4) Hybrid SOM-FOA model integrating Self-Organizing Maps with Firefly Optimization. Each method targets the core challenges of energy depletion, redundant transmission, and limited network lifetime. MATLAB-based simulations validate that all four methods achieve significant improvements over conventional techniques such as LEACH, PEGASIS, ACO, CS, and TREEPSI, with TDAC reducing energy consumption by up to 35%, improving end-to-end delay by up to 45%, and extending network lifetime by up to 17%.

Keywords: *Wireless Sensor Networks, Data Aggregation, Ant Colony Optimization, Cuckoo Search, IoT, Energy Efficiency, Cluster Head Selection, UAV Relay, SOM, Firefly Algorithm.*

1. Introduction

The Internet of Things (IoT) relies on wireless sensor networks (WSNs) to collect data from physical sensors across diverse domains, including military surveillance, building automation, smart agriculture, environmental monitoring, and healthcare. Sensor nodes are typically small, battery-powered devices deployed in large numbers, making energy efficiency and network longevity critical design parameters.

Data aggregation has emerged as a key technique for reducing redundant transmissions and extending device lifetimes. In an aggregation scheme, intermediate nodes combine data from neighboring devices before forwarding it toward the base station, reducing total transmissions and conserving energy. Despite extensive research, existing methods remain challenged by unbalanced energy consumption, hot-spot formation near the base station, slow convergence in metaheuristic approaches, and poor scalability to large, heterogeneous networks.

This paper reviews and compares four distinct data aggregation strategies — TDAC (hybrid ACO-CS tree-based), heterogeneous energy-aware aggregation, Quad Tree UAV-relay clustering, and SOM-FOA cluster-based aggregation — each addressing different aspects of the core WSN problem. The analysis covers algorithmic design, simulation methodology, performance metrics, and comparative results against established benchmarks.

2. Related Work and Research Gaps

Numerous aggregation strategies have been proposed for WSNs. Clustering protocols such as LEACH randomly select cluster heads, leading to inefficient energy use due to the assumption of homogeneous energy. PEGASIS reduces transmissions via chain structures but introduces significant delay. HEED improves cluster head selection using residual energy but does not fully address load balancing in heterogeneous environments.

Tree-based aggregation methods offer structured routing paths but suffer from uneven energy distribution. Metaheuristic approaches, including ACO, GA, and PSO, optimize routing paths but often converge slowly or are computationally expensive. Security-aware schemes such as SDAACA and hybrid QoS-aware approaches like QADA provide notable improvements but remain unbalanced regarding scalability and simultaneous multi-metric optimization.

Key research gaps identified across the literature include: (i) lack of dynamic role assignment adapting to changing energy conditions; (ii) absence of simultaneous multi-benchmark optimization for energy level, collision prevention, and neighbour energy; (iii) inadequate handling of large-scale heterogeneous deployments and energy hole prevention; and (iv) limited integration of security, mobility, and adaptive mechanisms in a unified framework.

3. Proposed Data Aggregation Methods

3.1 TDAC: Hybrid ACO-Cuckoo Search Tree-Based Aggregation

TDAC (Tree-based Data Aggregation using ACO and CS) constructs an optimal WSN aggregation tree by hybridizing Ant Colony Optimization with the Cuckoo Search algorithm. Each network node is assigned both an ant agent and a cuckoo agent. Ants explore routes to the base station using pheromone trails (τij) and heuristic distance matrices (ηij). CS employs Lévy flight-based random walks to locally refine ant solutions, replacing inferior solutions with superior cuckoo-generated ones. The network is modeled as a graph G = (V, E) with nodes randomly distributed over a 100×100 area. The fitness function is $F = a_1 \cdot N + a_2 \cdot D + a_3 \cdot E$, balancing network lifetime (N), delay (D), and energy (E). Pheromone trails are dynamically updated using evaporation rate ρ and reinforcement factor Q. Total computational complexity is $O(tmn^2) + O(tn)$.

3.2 Energy-Aware Aggregation for Heterogeneous Nodes

This method addresses the practical reality that sensor nodes exhibit different initial energy levels due to manufacturing variations and environmental conditions. Nodes are classified into high, medium, and low energy categories. High-energy nodes serve as aggregation points; medium-energy nodes act as forwarders; low-energy nodes perform only sensing. The aggregation tree is constructed dynamically — in each iteration, node energies are updated, and roles are reassigned based on residual energy, ensuring balanced depletion and preventing energy holes. The energy model follows: $E\_tx(k,d) = E\_elec \cdot k + E\_amp \cdot k \cdot d^2$ and $E\_rx(k) = E\_elec \cdot k$.

3.3 Quad Tree Clustering with Edge-Assisted UAV Relay

This framework integrates four sequential mechanisms: (1) Quadtree-based clustering organizes the network by node density, subdividing dense regions and simplifying sparse ones, with cluster heads selected using weighted criteria including residual energy, node degree, centrality, connectivity, and link stability; (2) Dynamic duty cycling using an Improved Unscented Kalman Filter schedules nodes among active, transmit, and sleep states based on residual energy, expected coverage rate, buffer factor, and historical behavior; (3) Twin-Agent Twin Delayed Deep Deterministic (TA-TD3) framework performs coverage hole detection and recovery, with separate agents for each task learning via reward-based feedback; (4) Cat and Mouse Optimization selects the optimal edge-assisted UAV relay node to minimize long-range communication burden and reduce end-to-end delay.

3.4 SOM-FOA Hybrid Cluster-Based Aggregation

This approach integrates Self-Organizing Maps (SOM) for adaptive clustering with the Firefly Optimization Algorithm (FOA) for cluster head selection. SOM groups sensor nodes based on location, residual energy, distance to base station, and SINR, preserving topological relationships.

FOA then selects optimal cluster heads using the fitness function: $f_i = \Sigma (W_1 \cdot E + W_2 \cdot S + W_3 \cdot (1-R) + W_4 \cdot (1-H))$, where E is residual energy, S is SINR, R is inter-node distance, and H is hopping count. Cluster heads collect, de-duplicate, aggregate, and forward data to the base station. The method is validated using the Intel Berkeley Research Lab dataset with 54 sensor nodes.

4. Simulation Setup and Results

4.1 Simulation Parameters

All methods were evaluated in MATLAB. Common parameters include network dimension 100×100 m, node count 5–500 (or 50–500), packet size 800 bits, E_{elec} = 50 nJ/bit, E_{fs} = 10 pJ/bit/m², E_{amp} = 0.0013 pJ/bit/m⁴, and pheromone evaporation rate $\rho = 0.05$. The TDAC simulation used MATLAB R2020b on a CPU Core i7 with 8 GB RAM; the SOM-FOA simulation used MATLAB 2024b with the Intel Berkeley dataset; and the Quad Tree framework used NS-3.26 with 100 nodes in a 500×500 m area.

4.2 TDAC Performance Results

TDAC was benchmarked against QADA, binary tree, SDAACA, ACO, CS, SAA, and TREEPSI across five metrics: total energy consumption, end-to-end delay, network lifetime, alive nodes, and communication overhead.

Metric	**TDAC Improvement**
Energy Consumption vs ACO	35% reduction
Energy Consumption vs CS	20% reduction
Energy Consumption vs SAA	13% reduction
End-to-End Delay vs Binary Tree	45% improvement
End-to-End Delay vs SDAACA	15% improvement
Network Lifetime vs TREEPSI	17% extension
Network Lifetime vs QADA	11% extension
Communication Overhead vs TREEPSI	11% reduction

Communication Overhead vs QADA	6% reduction

Benchmark validation across seven standard optimization functions (Powell Sum, Rosenbrock, Ackley, Weierstrass, Modified Schwefel, Sphere, Xin-She Yang) confirmed TDAC outperforms GA, PSO, BBO, and ABC, establishing it as a superior optimizer beyond WSN-specific applications.

4.3 Heterogeneous Energy-Aware Method Results

Compared against LEACH, PEGASIS, and HEED:

Nodes	**Proposed (J)**	**LEACH (J)**	**PEGASIS (J)**
50	0.12	0.22	0.18
100	0.20	0.36	0.29
200	0.35	0.60	0.48
500	0.80	1.50	1.18

Network lifetime (rounds) with the proposed method reached 9,000 at 500 nodes versus 6,000 (LEACH) and 7,200 (PEGASIS). End-to-end delay was reduced by up to 41% over LEACH at 500 nodes, and communication overhead was consistently 40–50% lower than LEACH across all node counts.

4.4 Quad Tree UAV Relay Results

Metric	**Proposed**	**Coherent**	**Repair**	**HWSN**
Energy Consumption (J)	18.4	42.5	45.0	55.0
Packet Delivery Ratio (%)	87.5	82.5	68.5	75.0
Network Lifetime (%)	95	90	82.5	86
Throughput (Mbps)	2.27	1.10	0.27	0.75

Delay (ms)	35.5	83.5	87.0	89.0
Coverage (%)	95.5	76.5	63.5	67.5

The integrated framework reduced energy consumption by 57% versus the coherent approach and achieved the highest throughput and coverage among all compared methods, validating the benefit of combining clustering, duty cycling, hole repair, and UAV relay selection.

4.5 SOM-FOA Results

Evaluated on the Intel Berkeley dataset with 54 nodes over 2000 simulation rounds, SOM-FOA achieved 15% improvement in network lifetime and 10% reduction in energy consumption versus individual SOM and standalone FOA, with improved throughput and load balancing due to the multi-parameter cluster head selection.

5. Comparative Analysis

Method	**Aggregation Type**	**Key Technique**	**Primary Strength**
TDAC	Tree-Based	Hybrid ACO + Cuckoo Search	Energy, delay, lifetime optimization
Heterogeneous EA	Tree-Based	Dynamic role assignment	Handles unequal initial energy
Quad Tree + UAV	Hierarchical Spatial	TA-TD3 + UAV relay	Coverage, lifetime, throughput
SOM-FOA	Cluster-Based	SOM clustering + FOA CH selection	Adaptive AI-driven clustering

All four methods address energy efficiency as a primary goal but differ in approach. TDAC excels in routing optimization for homogeneous networks. The heterogeneous energy-aware method is most suitable for practical deployments with unequal battery levels. The Quad Tree UAV framework delivers the most comprehensive multi-stage solution with the highest coverage and throughput. SOM-FOA is best suited for AI-assisted, data-driven IoT deployments with known topological constraints.

6. Conclusion and Future Work

This paper presented a comparative analysis of four advanced data aggregation methods for WSN-based IoT systems. Each method makes distinct contributions: TDAC hybridizes metaheuristic search for superior routing; the heterogeneous energy-aware scheme adapts dynamically to real-world node variability; the Quad Tree UAV framework integrates clustering, scheduling, coverage repair, and relay selection in a unified pipeline; and SOM-FOA combines unsupervised learning with bio-inspired optimization for intelligent cluster management.

Simulation results collectively confirm that all four approaches outperform conventional baselines, including LEACH, ACO, TREEPSI, and QADA, across metrics of energy consumption, network lifetime, delay, and communication overhead. The proposed methods together represent a comprehensive toolkit for designing energy-efficient, scalable, and robust WSN deployments.

Future research directions include: incorporation of node mobility and QoS awareness; integration of lightweight security mechanisms balancing data protection with energy efficiency; adaptive real-time reconfiguration for dynamic network topologies; and testing across larger-scale, heterogeneous real-world IoT environments. Comprehensive overhead evaluation and deep-learning-enhanced scheduling also remain promising avenues.

References

[1] Banafaa M et al. 6G mobile communication technology. Alex Eng J, 64:245–274, 2023.

[2] Zhang L. Tree-based energy-aware data aggregation in IoT using firefly optimization. Multiscale Multidisciplinary Modeling, 2023.

[3] Saeed N et al. Incorporating big data and IoT in intelligent ecosystems. Multimedia Tools and Applications, 83(7):20699–20741, 2024.

[4] Shoeibi M et al. Six-dimensional chimp optimization for reconfigurable intelligent surface energy harvesting in IoT. Future Internet, 16(12):460, 2024.

[5] Begum BA, Nandury SV. Data aggregation protocols for WSN and IoT — A comprehensive survey. J King Saud Univ, 35(2):651–681, 2023.

[6] Amodu OA et al. Age of Information minimization in UAV-aided WSN data collection. J Netw Comput Appl, 216:103652, 2023.

[7] Li Z et al. Lifetime balanced data aggregation for the Internet of Things. Comput Electr Eng, 58:244–264, 2017.

[8] Dehkordi SA et al. A survey on data aggregation in IoT sensor networks. Wireless Netw, 26(2):1243–1263, 2020.

[9] Razaque A, Rizvi SS. Secure data aggregation using access control in WSNs. Computers Security, 70:532–545, 2017.

[10] Rahman H et al. QoS-aware hybrid data aggregation for IoT. Ann Telecommun, 73(7–8):475–486, 2018.

[11] Hasheminejad E, Barati H. Reliable tree-based data aggregation in WSNs. Peer-to-Peer Networking and Applications, 14(2):873–887, 2021.

[12] Babukarthik R et al. Hybrid ACO and Cuckoo Search for job scheduling. Advances in Computing and Information Technology. Springer, 2013.

[13] Houssein EH et al. Lévy flight distribution: a new metaheuristic for engineering optimization. Eng Appl Artif Intell, 94:103731, 2020.

[14] Ullah I, Youn HY. Data aggregation based on self-organized map for WSN. J Supercomputing, 2019.

[15] Mosavvar I, Ghaffari A. Data aggregation in WSNs using firefly algorithm. Wireless Personal Communications, 2018.

[16] Hassan et al. Improved energy-efficient clustering protocol for WSN-IoT. IEEE Access, 8:200500–200517, 2020.

[17] Liao W-H et al. Data aggregation in WSNs using ant colony algorithm. J Netw Comput Appl, 31(4):387–401, 2008.

[18] Dhivya M, Sundarambal M. Cuckoo search for data gathering in WSNs. Int J Mobile Commun, 9(6):642–656, 2011.

[19] Satapathy SS, Sarma N. TREEPSI: tree-based energy efficient protocol for sensor information. IFIP WOCN, 2006.

[20] Finkel RA, Bentley JL. Quad trees: A data structure for retrieval on composite keys. Acta Informatica, 4:1–9, 1974.