Adaptive Crow Search Algorithm for Hierarchical Clustering in Internet of Things-Enabled Wireless Sensor Networks

Lingwei WANG, Hua WANG*
Guangdong University of Science and Technology, Dongguan, Guangdong, 523083, China

***Abstract*—The Internet of Things (IoT) relies on efficient Wireless Sensor Networks (WSNs) for data collection and transmission in various applications, including smart cities, industrial automation, and environmental monitoring. Clustering is a fundamental technique for structuring WSNs hierarchically, enabling load balancing, reducing energy consumption, and extending network lifespan. However, clustering optimization in WSNs is an NP-hard problem, necessitating heuristic and metaheuristic approaches. This study introduces an Adaptive Crow Search Algorithm (A-CSA) for clustering in IoT-enabled WSNs, addressing the inherent limitations of the standard CSA, such as premature convergence and local optima entrapment. The proposed A-CSA incorporates three key enhancements: 1) a dynamic awareness probability to improve global search efficiency during initial population selection, 2) a systematic leader selection mechanism to enhance exploitation and avoid random selection bias, and 3) an adaptive local search strategy to refine cluster formation. Performance evaluations conducted under varying network configurations, including node density, network size, and base station positioning, demonstrate that A-CSA outperforms existing clustering approaches in terms of energy efficiency, network longevity, and data transmission reliability. The results highlight the potential of A-CSA as a robust optimization technique for clustering in IoT-driven WSN environments.**

***Keywords*—Internet of things; wireless sensor networks; clustering; energy efficiency; optimization**

I. Introduction

A. Background and Motivation

The rapid expansion of the Internet of Things (IoT) has driven the widespread deployment of Wireless Sensor Networks (WSNs) for real-time data collection and communication in diverse applications [1]. Recently, WSNs have been extensively used in different sectors since they are affordable and easy to set up [2]. WSN comprises several diminutive Sensor Nodes (SNs) with restricted power supplies placed in surveillance regions to gather environmental information, including humidity, temperature, pressure, vibrations, and other characteristics [3]. The data is collected by combining many wireless connections and sent to a central station for assessment and processing [4].

Data transmission is the primary source of energy consumption in WSNs. SNs dissipate significant amounts of energy as packets of data are sent from one end of the network to the other. SNs are at risk of running out of energy while sending data packets owing to their low battery capacity. This issue can lead to premature network failure [5]. If not resolved, it poses a significant danger to the longevity of these networks. Reducing the size of data packets exchanged between SNs and finding the best routes for transmitting these packets are effective methods of managing energy usage in WSNs, resulting in enhanced and prolonged network functionality [6].

B. Research Problem and Challenges

WSNs trace their roots back to the "Tropical Tree" sensor system employed in the Vietnam War. Shahraki, et al. [7], introduced a novel algorithm that synergizes clustering strategies with compressed sensing, providing formal proofs for optimal cluster size, Cluster Head (CH) distribution, and inter-layer relationships. This approach effectively mitigates "hot issues" and curbs energy consumption arising from frequent CH role rotations. Recent advancements in low-power communication and signal processing technologies have facilitated widespread WSN deployment.

While cluster formation and CH selection are fundamental to WSNs, traditional protocols like LEACH and its centralized variant face limitations due to increasing network demands [8]. As WSNs are primarily battery-powered with dynamic topologies and unfixed node IDs, specialized routing protocols are imperative [9]. Clustering algorithms divide the network into manageable clusters. Network longevity relies on efficient energy management, and routing protocols have evolved accordingly [10].

In parallel with these developments, cross-disciplinary advancements have further enriched the analytical frameworks supporting WSNs. For instance, machine learning techniques have been employed to model and predict system behavior under complex environmental and economic variables, offering insights into adaptive strategies for WSN optimization [11]. Similarly, advanced simulation models, such as those used for weak rock mass behavior, demonstrate the importance of integrating environmental variability into system design and reliability assessments [12].

C. Authors' Contribution

This paper introduces a novel clustering protocol based on the Adaptive Crow Search Algorithm (A-CSA). This study seeks to answer the central question: How can the basic CSA be enhanced through dynamic parameter tuning, structured leader selection, and adaptive local search to improve energy-efficient clustering and prolong network lifetime in IoT-enabled WSNs?

To alleviate the energy consumption burden on Cluster Heads (CHs), we propose a relay node-assisted approach. Each CH is assigned a dedicated Relay Node (RN), eliminating the need for CHs to select next-hop nodes and reducing channel contention. RN selection is based on available energy, distance to the Base Station (BS), and proximity to the CH. To optimize CH and RN selection, we formulate a bi-objective fitness function that considers node position and residual energy. Given the NP-hard nature of this problem, A-CSA is employed to derive optimal solutions efficiently.

The body of the paper is formatted in the following way. A review of related work in WSN clustering and metaheuristic algorithms is presented in Section II, summarizing key developments and challenges. System model, network architecture, and energy considerations are explained in Section III. Section IV describes the clustering approach in detail. Section V discusses the enhanced CS algorithm for cluster node updating. Simulated results are presented in Section VI. The paper concludes with a summary and suggestions for further research in Section VII.

II. Literature Review

An overview of clustering and optimization protocols in WSNs is provided in this section. A comparison of metaheuristic and hybrid optimization techniques is presented in Table I, highlighting the key contributions of different approaches, algorithms, and performance metrics.

Chandirasekaran and Jayabarathi [13], developed a novel WSN protocol that leverages the Cat Swarm Optimization (CSO) algorithm for real-time clustering. Their approach minimizes distances within the cluster and optimizes energy distribution. By considering received signal strength, remaining battery voltage, and distance within the cluster, CSO effectively selects CHs. Performance evaluations against LEACH-C and PSO showed significantly improved battery life compared to traditional methods.

Mehta and Saxena [14], suggested a novel clustering and routing approach for WSNs using Sailfish Optimizer (SFO) to improve energy efficiency. A multi-objective fitness function guides the selection of CHs, prioritizing energy conservation and minimizing node failures. SFO determines optimal data transmission paths to the sink node. A comparative analysis of GWO, GA, ALO, and PSO shows superior performance for energy consumption and network lifespan, with corresponding improvements of 21% and 24% over GWO, respectively.

Nabavi, et al. [15], offered a novel hybrid approach to optimizing WSNs, combining genetic and Gravitational Search (GS) algorithms. The genetic algorithm was employed for CH selection, aiming to minimize intra-cluster distances and energy consumption, while GS was utilized for efficient routing between CHs and the sink node. The proposed technique demonstrated superior efficiency regarding energy efficiency, network throughput, and data delivery rate compared to existing techniques.

TABLE I STUDIES ON CLUSTERING AND OPTIMIZATION STRATEGIES IN WSNS

Study	Algorithm(s) used	Contribution	Performance metrics	Shortcoming
[13]	Cat swarm optimization	Real-time clustering, minimizing intra-cluster distances, and optimizing energy distribution	Improved battery life	Premature convergence, lacks adaptive search mechanism
[14]	Sailfish optimizer	Multi-objective CH selection and optimal data transmission paths	Energy consumption and network lifespan	High computational complexity, lacks relay node optimization
[15]	Genetic and gravitational search algorithms	Hybrid approach for CH selection and efficient routing	Energy efficiency, network throughput, and data delivery rate	Increased computational overhead, no adaptive clustering strategy
[16]	Ant colony optimization and butterfly optimization algorithms	Energy-efficient clustering with mobile sink option to mitigate hotspot problem	Residual energy, throughput, and alive nodes	High convergence time, lacks adaptive exploration-exploitation balance
[17]	Differential evolution and sparrow search algorithms	Hybrid approach for energy-efficient CH selection	Network lifespan, residual energy, and throughput	Susceptible to local optima, lacks adaptive parameter tuning
[18]	Levy chaotic particle swarm optimization	Enhanced convergence and search space in cluster routing, focusing on realistic industrial conditions	Energy usage and network longevity	Lacks adaptive control over search balance
[19]	Capuchin search algorithm and fuzzy logic	Energy-efficient data aggregation with multi-phase cluster formation and routing	Energy usage, delay, packet delivery rate, and network lifespan	Complexity in multi-phase processing, no dynamic relay node selection
[20]	Fuzzy logic and quantum annealing	On-demand clustering and energy-efficient routing to extend WSN durability	Network lifetime, energy consumption, and throughput	Computational complexity and dependency on predefined thresholds

Amutha, et al. [16], developed a novel energy-efficient clustering algorithm for WSNs that combines Ant Colony Optimization (ACO) and Butterfly Optimization (BO) algorithms. BO is employed for optimal CH determination, while ACO is used for energy-aware routing. To further extend network lifespan, two variants are introduced: HBACS with a fixed sink and HBACM with a mobile sink. The latter mitigates the hotspot problem by eliminating multi-hop communication between CHs and the sink. Simulation results with NS2 show significant improvements in residual energy, active nodes, and throughput for both variants compared to conventional algorithms.

Kathiroli and Selvadurai [17], suggested a hybrid optimization approach combining Differential Evolution (DE) and Sparrow Search Algorithm (SSA) to enhance energy efficiency in WSNs through optimized CH selection. Leveraging SSA's global search capability and DE's local search potential, the proposed algorithm effectively extends the network lifetime. Evaluation metrics include residual energy, network lifetime, throughput. Compared to existing methods, the hybrid SSA-DE approach demonstrated improved residual energy and throughput, highlighting its effectiveness in selecting optimal CHs.

Luo, et al. [18], developed an enhanced Levy Chaotic Particle Swarm Optimization-based Cluster Routing Protocol (LCPSO-CRP) for extending WSN lifetime. By introducing a chaotic optimization methodology, the protocol significantly accelerates convergence and expands the range of possible solutions. This innovative strategy, in accordance with BS distance, cluster-to-cluster distance, and node energy levels, outperforms traditional schemes like DEEC, LEACH, LEACH-kmeans, and LEACH-C. Extensive simulations under realistic industrial conditions demonstrate a minimum 22.9% reduction in energy consumption and a 13.9% extension of network longevity for LCPSO-CRP.

Mohseni, et al. [19], proposed a novel energy-efficient data aggregation routing mechanism for WSNs that couples the Capuchin Search Algorithm (CapSA) and fuzzy logic operators. This multi-phase approach comprises two primary steps: cluster creation and internal/external routing. Simulations using MATLAB validate the superiority of the suggested design regarding energy usage, delay, packet delivery rate, and network lifespan compared to existing approaches.

Wang, et al. [20], designed a hybrid routing and clustering protocol (FQA) combining fuzzy logic and quantum annealing to maximize WSN durability and decrease energy usage. Fuzzy logic is employed for intelligent CH selection, while quantum annealing optimizes data routing to the BS. An energy threshold mechanism is incorporated to expedite the process. Unlike traditional periodic clustering, FQA adopts an on-demand approach to reduce computational overhead. Comparative analysis against FC-RBAT, OAFS-IMFO, BOA-ACO, and FRNSEER consistently demonstrates FQA's better performance by demonstrating better network lifespan, data transfer rate, and energy usage across various scenarios.

Existing clustering and optimization techniques for WSNs have demonstrated notable improvements in energy efficiency, network lifespan, and data transmission reliability. However, several challenges remain unaddressed. Many approaches, including those based on CSO, SFO, and hybrid metaheuristic techniques, suffer from premature convergence, leading to suboptimal CH selection and inefficient energy distribution. Additionally, while hybrid methods enhance solution quality by leveraging multiple algorithms, they often introduce excessive computational overhead, making them less practical for real-time WSN applications.

Some techniques, such as those using fuzzy logic or chaotic optimization, improve network longevity but lack adaptability to dynamic network conditions. Furthermore, relay node selection remains an overlooked aspect, with most studies focusing solely on CH selection without optimizing multi-hop communication. To address these limitations, our proposed algorithm introduces dynamic awareness probability, systematic leader selection, and adaptive local search, ensuring balanced exploration and exploitation while enhancing clustering efficiency and energy management in IoT-driven WSNs.

III. System Model

For real-time environmental monitoring, a WSN of N SNs is considered. As shown in Fig. 1, each SN is equipped with a microcontroller unit, a communication unit, and a power management unit. SNs exhibit identical capabilities to operate in sensing or communication modes, collecting environmental data or transmitting information, respectively. Each SN possesses a data link to handle all data traffic. Nodes are spatially indexed. Static network topology is assumed to be consistent with typical WSN deployments. Table II summarizes the symbols and their definitions used in whole body of the study.

SNs are initialized with equal energy levels, creating a homogeneous network. It is impossible to replace the battery, simulating unattended operation. Environmental data are collected at fixed intervals and transmitted periodically. SNs can adjust transmission power across multiple levels based on recipient distance. Bidirectional communication links exist between nodes, with distance estimation relying solely on received signal strength. Exploiting data correlation, CHs compress gathered data into fixed-length packets. The BS maintains a continuous power supply.

This study adopts a simplified communication energy consumption model incorporating path loss effects, as illustrated in Fig. 1. The propagation channel is characterized by either free space (inverse square law) or multipath fading (inverse fourth power law) attenuation, contingent upon transmitter-receiver distance. Power control mitigates these losses. For distances below a threshold, d_0, free space propagation is assumed; otherwise, the multipath model prevails.

TABLE II SYMBOLS AND DEFINITIONS

Symbol	Definition	Symbol	Definition
N	Number of sensor nodes	d_0	Threshold distance for free-space vs. multipath propagation
K	Data packet size	E_{TX}	Transmission energy consumption
d	Transmission distance between nodes	E_{elec}	Circuit energy dissipation per bit
E_{RX}	Reception energy consumption	E_{fs}	Free-space model amplifier energy
E_{mp}	Multipath fading model amplifier energy	k	Number of active (operational) nodes at a given time
ξ	Predefined threshold for the proportion of dead nodes	FND	Time until the first sensor node depletes its energy
PND	Time until a specified proportion of nodes are dead	$\widetilde{CH}$	Set of non-CH nodes
n	Number of clusters formed	R^{CH}_{energy}	CH energy ratio relative to non-CH nodes
F_{CH}	Fitness function for CH selection	$R^{CH}_{location}$	CH location ratio relative to non-CH nodes
$E^{res}_{CH}(j)$	Residual energy of CH node j	a	Weighting factor for fitness components
$\bar{E}_{CH}$	Average residual energy of CHs	$\bar{D}_{CH}$	Average distance of CHs from the BS
$\bar{L}_{CN}$	Average distance of CNs from BS and CHs	F_{RN}	Fitness function for relay node selection
$x_{i,t}$	Position of crow i at time t	$m_{i,t}$	Memory position of crow i at time t
fl	Flight length factor	K_t	Dynamic awareness probability at generation t
K_{max}	Maximum awareness probability	K_{min}	Minimum awareness probability
t_{max}	Maximum number of generations	AP	Awareness probability
D_{thr}	Threshold distance for adaptive flight length	$gb_{j,t}$	Best position of crow j at time t

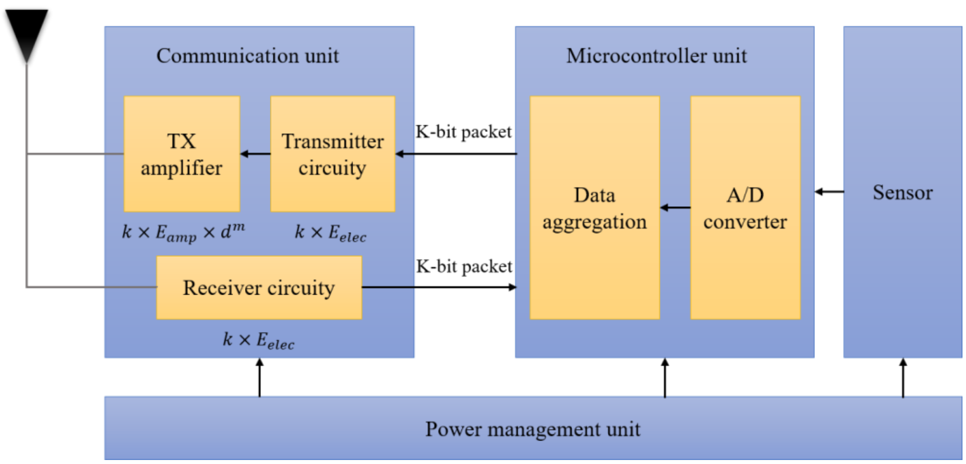

Fig. 1. Components and configuration of SNs in WSNs.

Eq. (1) quantifies the energy expended in relaying a *K-bit* packet across a distance d. The components of this energy consumption model include transmission energy (E_{TX}), reception energy (E_{fs}), link distance (d), circuit-level energy dissipation (E_{elec}), amplifier model parameters (E_{fs}, E_{mp}), data packet length (k), and the transmission distance threshold (d_0) defined in Eq. (2). Eq. (3) determines packet reception energy consumption. E_{elec} encompasses the energy consumed by transmitter or receiver circuitry, influenced by factors such as signal spreading, filtering, modulation, and channel coding.

$$E_{TX}(k,d) = \begin{cases} k \times E_{elec} + k \times E_{fs} \times d^2 & if\ d \leq d_0 \\ k \times E_{elec} + k \times E_{mp} \times d^4 & if\ d > d_0 \end{cases} \quad (1)$$

$$d_0 = \sqrt{\frac{E_{fs}}{E_{mp}}} \quad (2)$$

$$E_{RX}(k) = k \times E_{elec} \quad (3)$$

The resilience of many WSN applications to node failures is evident, particularly in high-density deployments where redundant sensing capabilities exist among neighboring nodes. Consequently, the duration before First Node Dead (FND) is an insufficient metric for comprehensive network lifetime assessment. A more robust indicator, the time until a predefined percentage of nodes dead (PND), is proposed for scenarios characterized by high node density. Eq. (4) defines network lifetime as the duration before the proportion of operational nodes reaches a specified threshold, ξ. Within this equation, N denotes sensor count and k signifies the number of active nodes.

$$T_N^k = T\left[\xi = \frac{k}{N}\right] \quad (4)$$

IV. CLUSTERING APPROACH

The proposed protocol categorizes nodes into CHs, Common Nodes (CNs), and RNs. Network configuration comprises repeated cluster formation and data transmission stages. During the setup phase, clusters, CHs, RNs, and routing paths to the BS are established. The data transmission phase involves data collection by CHs from cluster members, relayed through RNs to the BS, following a predefined scheme. Fig. 2 illustrates the network architecture.

Given N randomly deployed SNs, n clusters are formed. The set of CHs is denoted as CH, while non-CH nodes are represented by $\widetilde{CH}$. CHs coordinate cluster operations, aggregate data, and communicate with RNs. CH selection considers node energy levels and locations, prioritizing nodes with more available energy and closer to the BS for balanced cluster formation. This optimization problem is formulated in Eq. (5).

$$F_{CH} = a \times R^{CH}_{energy} + (1 - a) \times R^{CH}_{location} \quad (5)$$

The fitness function, F_{CH}, comprises two components: R^{CH}_{energy} and $R^{CH}_{location}$, weighted by α. R^{CH}_{energy}, calculated in Eq. (6), represents CH energy relative to non-CH energy, favoring energy-rich nodes as CHs. $R^{CH}_{location}$, determined by Eq. (7), measures the relative distance of CHs and non-CHs to the BS, promoting CHs closer to the BS for improved energy efficiency.

$$R^{CH}_{energy} = \frac{\bar{E}_{CH}}{\bar{E}_{\widetilde{CH}}} = \frac{\sum_{\forall node_j \in CH} E^{res}_{CH}(j)/|CH|}{\sum_{\forall node_j \in \widetilde{CH}} E^{res}_{\widetilde{CH}}(j)/|\widetilde{CH}|} \quad (6)$$

$$R^{CH}_{location} = \frac{\bar{D}_{\widetilde{CH}}}{\bar{D}_{CH}} = \frac{\sum_{\forall node_j \in \widetilde{CH}} d(node_i, BS)/|\widetilde{CH}|}{\sum_{\forall node_j \in CH} d(node_i, BS)/|CH|} \quad (7)$$

Practical WSNs employ battery-powered nodes, with residual energy indicated by the battery voltage. Nodes with higher energy and proximity to the BS exhibit a higher likelihood of becoming CHs. Given the NP-hard nature of this problem, an improved CS algorithm is proposed for the solution, as detailed in the subsequent section.

To mitigate excessive energy consumption among CHs, RNs are introduced to share the data transmission burden. RN selection depends on two primary criteria: superior energy levels relative to CNs and optimal spatial positioning between the CH and the BS to minimize energy-intensive transmissions. Unlike conventional approaches, our protocol assigns a dedicated RN to each CH, reducing communication overhead between these entities.

The RN selection process is guided by a fitness function (Eq. (8)) comprising two components: R^{EN}_{energy} and $R^{RN}_{location}$. R^{EN}_{energy}, calculated in Eq. (9), represents the ratio of average RN energy to average CN energy, prioritizing energy-rich nodes for RN roles. $R^{RN}_{location}$, defined in Eq. (10), evaluates the relative position of a potential RN to its corresponding CH and the BS, aiming to minimize transmission distances.

$$F_{RN} = \beta \times R^{RN}_{energy} + (1 - \beta) \times R^{RN}_{location} \quad (8)$$

$$R^{EN}_{energy} = \frac{\bar{E}_{RN}}{\bar{E}_{CN}} = \frac{\sum_{\forall node_z \in RN} E^{res}_{RN}(z)/|RN|}{\sum_{\forall node_k \in CN} E^{res}_{CN}(k)/|CN|} \quad (9)$$

$$R^{RN}_{location} = \frac{\bar{L}_{CN}}{\bar{L}_{RN}} = \frac{\sum_{\forall node_k \in CN}\{d(node_k, BS) + d(node_k, CH_j)\}/|CN|}{\sum_{\forall node_z \in RN}\{d(RN_z, BS) + d(RN_z, CH_j)\}/|RN|} \quad (10)$$

By maximizing both R^{EN}_{energy} and $R^{RN}_{location}$, the protocol ensures the selection of energy-efficient and strategically positioned RNs, thereby enhancing overall network performance. Similar to other protocols, CH and RN selection is centralized at the BS. SNs are given an identifier (ID) based on their location. The clustering process commences with nodes broadcasting residual energy and location information via *Node-MSG* messages. The BS selects CHs and disseminates their IDs through a broadcast message. Subsequently, CHs introduce themselves to the network using *CH-ADV* messages. A similar process is followed for RN selection and announcement (*RN-ADV* messages).

CNs determine their cluster membership by selecting the CH requiring the least transmission energy based on received *CH-ADV* messages. Upon cluster selection, nodes notify the CHs via *JOIN-REQ* messages. The CH functions as a control center, establishing a TDMA scheduler and disseminating it through *SCHEDULE-MSG* messages. This synchronization mechanism reduces energy usage and enhances spectral efficiency by enabling nodes to power down their radios during idle periods. The data transmission phase adheres to the TDMA schedule, with CNs sending data to their CHs. The CH collects data and relays it to the RN, which ultimately transmits aggregated data to the BS.

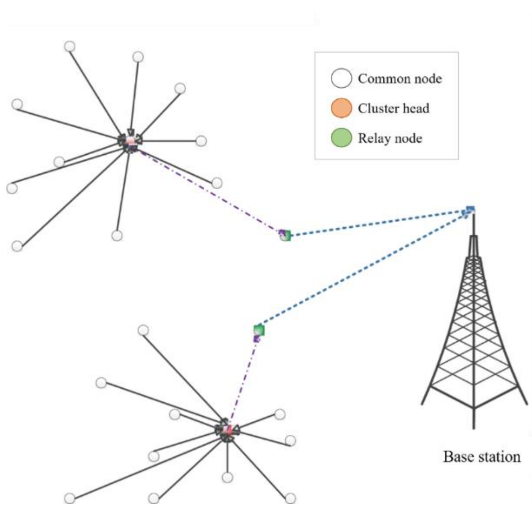

Fig. 2. Network architecture of the proposed protocol.

V. Proposed Algorithm for Cluster Node Updating

The conventional CS algorithm mimics the intelligent behavior of crows to solve optimization problems [21]. It involves the following steps:

- Initialization: Define the optimization problem, set parameters, and initialize crow positions randomly within defined bounds.
- Memory update: Store the initial position of each crow.

$$Crows_{i,t} = \begin{bmatrix} x_i^1 & \cdots & x_i^d \\ \vdots & \dots & \vdots \\ x_N^1 & \cdots & x_N^d \end{bmatrix} \tag{11}$$

$$CrowsMemory_{i,t} = \begin{bmatrix} m_i^1 & \cdots & m_i^d \\ \vdots & \dots & \vdots \\ m_N^1 & \cdots & m_N^d \end{bmatrix} \tag{12}$$

- Position update: If crow *j* is unaware of crow *i*, update crow *i*'s position using Eq. (13) and a local or global search based on *fl*. If crow *j* is aware of crow *i*, update crow *i*'s position randomly.

$$x_{i,t+1} = x_{i,t} + r_i \times fl \times (m_{j,t} - x_{i,t}) \tag{13}$$

$$x_{i,t+1} = a\ random\ position \tag{14}$$

- Evaluation: Evaluate the fitness of new positions and update crow memories with better solutions.
- Termination: Steps 2-4 are repeated until generation limit is met.

The Modified Crow Search (MCS) Algorithm is an improvement over the conventional CS. It introduces two key modifications:

Dynamic awareness probability: Instead of a fixed *AP*, the MCS algorithm uses a parameter *K* that dynamically adjusts the probability of a crow being aware of another crow's following. This promotes exploration in early generations and exploitation in later generations.

$$K_t = round\left(K_{max} - \frac{K_{max} - K_{min}}{t_{max}} \times t\right) \tag{15}$$

Adaptive flight length: The MCS algorithm calculates the flight length (*fl*) based on the distance between crows ($D_{i,j}$). This allows for a more focused search in promising regions.

$$fl_{i,t} = \begin{cases} 2 & if\ D_{i,j} > D_{thr} \\ fl_{thr} & if\ D_{i,j} \le D_{thr} \end{cases} \tag{16}$$

The conventional CS algorithm employs a repetitive optimization process that hinges on exploration and exploitation, modulated by the parameter *AP*, typically set to 0.1. This configuration predominantly biases the algorithm towards exploitation, often at the expense of exploration across all generations. As a consequence, the CS algorithm is susceptible to local optima and exhibits strong dependence on the initial population. To mitigate these limitations, this paper introduces a novel approach that incorporates a dynamically adjusted *AP*, correlated with the generation count, and employs two innovative equations to enhance exploration and exploitation.

$$AP = AP_{max} + \frac{AP_{max} - AP_{min}}{t_{max}} \times t \tag{17}$$

Similar to the standard CS, step 1 of A-CSA involves problem definition and parameter initialization. However, A-CSA introduces additional parameters: AP_{max}, AP_{min}, and *FAR* (Flight Awareness Ratio). AP_{max} and AP_{min} are pivotal for the dynamic *AP* mechanism.

Consistent with the conventional CS, A-CSA employs Eq. (11) and Eq. (12) to determine crew group size and initialize crow positions. Subsequently, the objective function evaluates the fitness of these initial positions.

A-CSA diverges significantly from the conventional CS in three key aspects. Firstly, A-CSA incorporates a dynamic *AP* that fluctuates with the generation count, as governed by Eq. (18). AP_{max} and AP_{min}, bounded between 0 and 1, control the exploration-exploitation balance. A larger *AP* promotes exploration, while a smaller *AP* favors exploitation. Optimal algorithm performance necessitates a judicious selection of *AP*.

$$AP_t = AP_{min} + \frac{AP_{max} - AP_{min}}{ln(t)+1} \tag{18}$$

Secondly, in contrast to the random selection of a crow to follow in the conventional CS (Eq. (13)), A-CSA employs a FAR-based mechanism to guide crow *i* towards the best crow *j* ($gb_{j,t}$) as defined by Eq. (19). $r_{i,t}^2$ and $r_{i,t}^3$ are random values within the [0, 1] interval, and *FAR* is a predefined constant between 0 and 1. This modification enhances exploitation compared to the standard CS. A *FAR* approaching 0 prioritizes memory-based best solutions, while a value closer to 1 resembles the random selection of the conventional CS. By tuning *FAR* appropriately, the algorithm can achieve a harmonious balance between exploration and exploitation, thereby improving convergence.

$$x_{i,t+1} = \begin{cases} x_{i,t} + r_{i,t}^2 \times fl \times (m_{j,t} - x_{j,t}) & if\ r_{i,t}^3 \le FAR \\ x_{i,t} + r_{i,t}^2 \times fl \times (gb_{j,t} - x_{j,t}) & else \end{cases} \tag{19}$$

Thirdly, A-CSA refines the exploration phase of the conventional CS. While the latter employs a random search within the lb and ub bounds when the random number exceeds *AP*, A-CSA introduces a localized search mechanism through Eq. (20) as the generation progresses. This strategy mitigates the diminishing returns of global search in later generations. $r_{i,t}^4$ and $r_{i,t}^5$ are random values within the [0, 1] range.

$$x_{i,t+1} = \begin{cases} 2x_{i,t} + \left(lb + r_{i,t}^5 \times (lb - ub)\right)/t & if\ r_{i,t}^4 \le 0.5 \\ a\ random\ position & else \end{cases} \tag{20}$$

The optimization process iteratively executes Steps 2-4 until the generation count reaches the predefined maximum (t_{max}), yielding the final solution. Algorithm 1 presents the pseudocode of A-CSA.

Algorithm 1 Pseudocode of A-CSA

Input:
Initialize control parameters: $AP_{max}, AP_{min}, FAR, fl, N, pd, t_{max}$
Randomly generate and store the initial positions of all crows in the solution space
Evaluate the initial fitness of each crow
Begin iteration:
Repeat until the maximum number of iterations t_{max} is reached:
Randomly select the crow positions to update
For each crow $i = 1$ to N:
  Compute the dynamic awareness probability AP_t
  Generate a random number $r^1 \in [0,1]$
  If $r^1 \geq AP_t$:
    Generate another random number $r^3 \in [0,1]$
    If $r^3 < FAR$:
      Update position using Equation 19a
    Else
      Update position using Equation 19b
  Else
    Generate another random number $r^4 \in [0,1]$
    If $r^4 \leq 0.5$:
      Update position using Equation 20
    Else
      Assign $x_{i,t+1}$ a random position within the bounds
End for
Evaluate the updated positions and compute fitness values
Update memory based on improved solutions
End repeat
Report the best-found solution and convergence results

The computational complexity of the proposed A-CSA algorithm primarily stems from the repeated evaluation of the fitness functions for CH and RN selection and the iterative update of crow positions. Let N represent the number of nodes, G the number of generations, and C the number of candidate crows. The per-generation complexity is $O(C{\cdot}N)$, resulting in an overall complexity of $O(G{\cdot}C{\cdot}N)$. As clustering and optimization are carried out at the base station, which is not resource-constrained, this overhead remains acceptable for practical deployments. Future enhancements may involve parallel implementations to reduce computation time further.

VI. Results and Discussion

This section presents a comprehensive investigation of the suggested method's effectiveness through simulation analyses. The protocol's efficacy in constructing energy-efficient routing hierarchies for WSNs is assessed, with a particular focus on applications demanding long network longevity and effective data aggregation, e.g., monitoring environments.

Key parameters, including network lifetime, node count, BS positioning, and network dimension, were considered in the comparative analysis of various routing protocols. MATLAB was employed for simulation modeling and programming, with average values derived from 20 simulation runs to enhance result reliability. Simulation variables are summarized in Table III.

Fig. 3 illustrates the convergence rate of the objective function's fitness value, demonstrating convergence within 50 iterations. Consequently, the algorithm's maximum iteration count was set to 50.

A comparative analysis with a CS-based protocol is presented in Fig. 4. Network lifetime metrics employed include FND, Last Node Dead (LND), and Half Number of Nodes Dead (HND). A-CSA significantly improved over the CS-based approach, with FND, HND, and LND extended by 98%, 101%, and 105%, respectively.

TABLE III SIMULATION VARIABLES

Feature	Variable	Value
Radio model	E_{DA}	*5nJ/bit/signal*
	d_0	*75m*
	E_{mp}	*0.0013pJ/bit/m⁴*
	E_{fs}	*10pJ/bit/m²*
	E_{elec}	*50nJ/bit*
Network	Initial energy of nodes	*2J*
	BS location	*(50,175)*
	Dimension	(0,0)~(100,100)

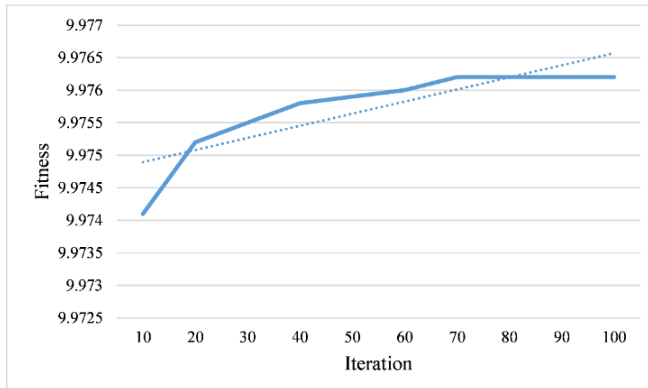

Fig. 3. Convergence rate.

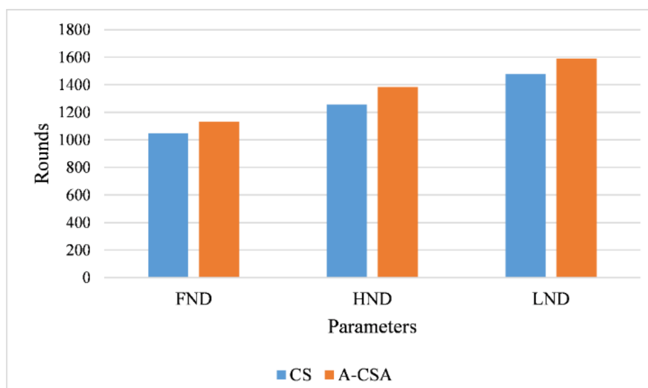

Fig. 4. Comparative analysis of network lifetime metrics.

To measure the lifetime efficiency of the suggested protocol, three distinct situations were considered varying in terms of node count, BS position, and network area size, as detailed in Table IV. The proposed protocol was benchmarked against SEECH, TCAC, and LEACH. Fig. 5, Fig. 6, and Fig. 7 visualize the count of alive nodes over time. A-CSA consistently outperformed the other three compared protocols.

TABLE IV EVALUATION SCENARIOS

Parameter	1st situation	2nd situation	3rd situation
Node count	100	500	1000
BS position	(50,175)	(50,200)	(100,300)
Dimension	(100,100)	(100,100)	(200,200)

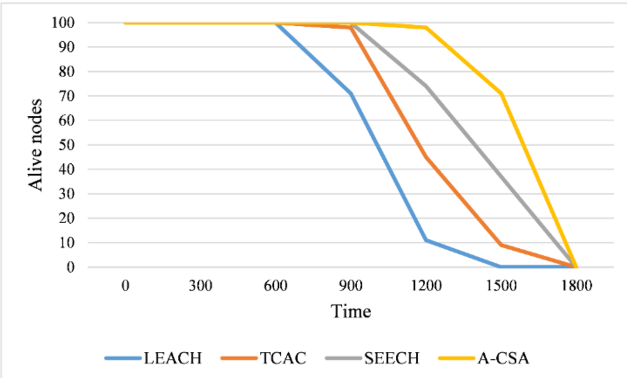

Fig. 5. Network lifetime comparison for first scenario.

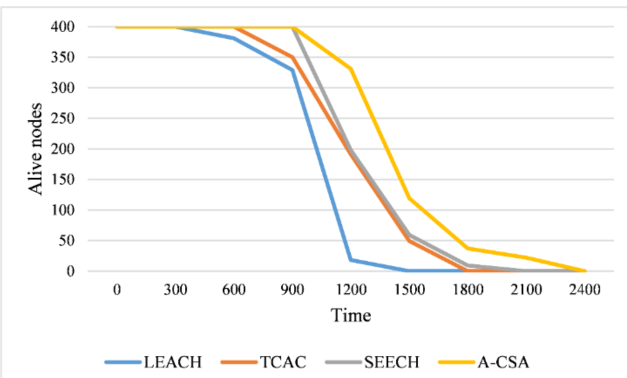

Fig. 6. Network lifetime comparison for second scenario.

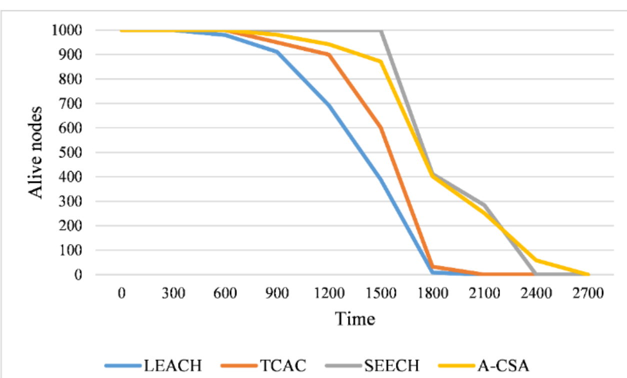

Fig. 7. Network lifetime comparison for third scenario.

Unlike existing models that rely on static or heuristic-based decisions, A-CSA dynamically adapts its search behavior over generations, resulting in better convergence and more energy-efficient clustering. Moreover, the integration of a bi-objective fitness function for both CH and RN selection provides an edge in maintaining balanced energy distribution, especially in dense and large-scale networks. These improvements position A-CSA as a competitive and scalable alternative to traditional and hybrid clustering approaches.

VII. CONCLUSION

Energy efficiency is a paramount challenge in WSNs. Clustering and routing strategies are commonly employed to address this issue. However, these problems are classified as NP-hard optimization problems, necessitating heuristic approaches. Swarm intelligence algorithms have emerged as promising candidates for obtaining near-optimal solutions to these complex challenges. This paper introduced a novel clustering protocol for WSNs that leverages RNs to alleviate the energy burden on CHs. An enhanced CS algorithm is proposed to optimize cluster formation, minimizing transmission distances and energy consumption. This approach effectively prolongs the network lifetime. Comprehensive simulations under diverse network conditions, including varying node densities, network areas, and BS positions, demonstrated the protocol's superior energy efficiency compared to existing clustering protocols. By balancing energy consumption among nodes and reducing overall energy expenditure, the proposed protocol significantly extends the network lifetime.

While the primary focus of this study was on energy efficiency and network longevity, critical factors in battery-constrained WSNs, we acknowledge that comprehensive validation should also consider metrics such as delay, throughput, and packet delivery ratio . Although our simulation results include node activity trends that indirectly reflect throughput, detailed quantitative evaluations of delay and PDR were not included in this version. Future work will incorporate these metrics to provide a more holistic assessment of the protocol's suitability for time-sensitive and high-reliability IoT applications. Additionally, to enhance the applicability of A-CSA in dynamic IoT environments, we aim to incorporate mobility models and traffic-aware mechanisms, enabling the protocol to adapt to node mobility, varying network topologies, and fluctuating traffic loads commonly encountered in real-world deployments.

FUNDING

This work was supported by the following projects: (1) the 2022 Dongguan Municipal Science and Technology Project for Social Development, titled "Research on Routing Protocol Algorithms Driven by Internet of Things Dedicated Frequency Networking" (Project No. 20221800902742); and (2) the 2024 Guangdong University of Science and Technology Key Research Project, titled "Optimization of Big Data Collection for Smart Agriculture Internet of Things from the Perspective of New-Form Productivity" (Project No. GKY-2024KYZDK-17).

REFERENCES

[1] M. Nassereddine and A. Khang, "Applications of Internet of Things (IoT) in smart cities," in Advanced IoT technologies and applications in the industry 4.0 digital economy: CRC Press, 2024, pp. 109-136.

[2] A. A. K. Majhi and S. Mohanty, "A Comprehensive Review on Internet of Things Applications in Power Systems," IEEE Internet of Things Journal, 2024.

[3] P. Kaur, K. Kaur, K. Singh, and S. Kim, "Early forest fire detection using a protocol for energy-efficient clustering with weighted-based optimization in wireless sensor networks," Applied Sciences, vol. 13, no. 5, p. 3048, 2023.

[4] B. Pourghebleh, N. Hekmati, Z. Davoudnia, and M. Sadeghi, "A roadmap towards energy - efficient data fusion methods in the Internet of Things," Concurrency and Computation: Practice and Experience, vol. 34, no. 15, p. e6959, 2022.

[5] K. Biswas, V. Muthukkumarasamy, M. J. M. Chowdhury, X.-W. Wu, and K. Singh, "A multipath routing protocol for secure energy efficient communication in Wireless Sensor Networks," Computer Networks, vol. 232, p. 109842, 2023.

[6] N. Meenakshi et al., "Efficient communication in wireless sensor networks using optimized energy efficient engroove leach clustering protocol," Tsinghua Science and Technology, vol. 29, no. 4, pp. 985-1001, 2024.

[7] A. Shahraki, A. Taherkordi, Ø. Haugen, and F. Eliassen, "Clustering objectives in wireless sensor networks: A survey and research direction analysis," Computer Networks, vol. 180, p. 107376, 2020.

[8] I. Daanoune, B. Abdennaceur, and A. Ballouk, "A comprehensive survey on LEACH-based clustering routing protocols in Wireless Sensor Networks," Ad Hoc Networks, vol. 114, p. 102409, 2021.

[9] T. Taami, S. Azizi, and R. Yarinezhad, "An efficient route selection mechanism based on network topology in battery-powered internet of things networks," Peer-to-Peer Networking and Applications, vol. 16, no. 1, pp. 450-465, 2023.

[10] Y. K. Sharma, G. Ahmed, and D. K. Saini, "Uneven clustering in wireless sensor networks: A comprehensive review," Computers and Electrical Engineering, vol. 120, p. 109844, 2024.

[11] M. B. Bagherabad, E. Rivandi, and M. J. Mehr, "Machine Learning for Analyzing Effects of Various Factors on Business Economic," Authorea Preprints, 2025, doi: https://doi.org/10.36227/techrxiv.174429010.09842200/v1.

[12] A. Azadi and M. Momayez, "Simulating a Weak Rock Mass by a Constitutive Model," Mining, vol. 5, no. 2, p. 23, 2025, doi: https://doi.org/10.3390/mining5020023.

[13] D. Chandirasekaran and T. Jayabarathi, "Cat swarm algorithm in wireless sensor networks for optimized cluster head selection: a real time approach," Cluster Computing, vol. 22, pp. 11351-11361, 2019.

[14] D. Mehta and S. Saxena, "MCH-EOR: Multi-objective cluster head based energy-aware optimized routing algorithm in wireless sensor networks," Sustainable Computing: Informatics and Systems, vol. 28, p. 100406, 2020.

[15] S. R. Nabavi, V. Ostovari Moghadam, M. Yahyaei Feriz Hendi, and A. Ghasemi, "Optimal selection of the cluster head in wireless sensor networks by combining the multiobjective genetic algorithm and the gravitational search algorithm," Journal of Sensors, vol. 2021, no. 1, p. 2292580, 2021.

[16] J. Amutha, S. Sharma, and S. K. Sharma, "An energy efficient cluster based hybrid optimization algorithm with static sink and mobile sink node for Wireless Sensor Networks," Expert Systems with Applications, vol. 203, p. 117334, 2022.

[17] P. Kathiroli and K. Selvadurai, "Energy efficient cluster head selection using improved Sparrow Search Algorithm in Wireless Sensor Networks," Journal of King Saud University-Computer and Information Sciences, vol. 34, no. 10, pp. 8564-8575, 2022.

[18] T. Luo, J. Xie, B. Zhang, Y. Zhang, C. Li, and J. Zhou, "An improved levy chaotic particle swarm optimization algorithm for energy-efficient cluster routing scheme in industrial wireless sensor networks," Expert Systems with Applications, p. 122780, 2023.

[19] M. Mohseni, F. Amirghafouri, and B. Pourghebleh, "CEDAR: A cluster-based energy-aware data aggregation routing protocol in the internet of things using capuchin search algorithm and fuzzy logic," Peer-to-Peer Networking and Applications, vol. 16, no. 1, pp. 189-209, 2023.

[20] H. Wang, K. Liu, C. Wang, and H. Hu, "Energy-Efficient, Cluster-Based Routing Protocol for Wireless Sensor Networks Using Fuzzy Logic and Quantum Annealing Algorithm," Sensors, vol. 24, no. 13, p. 4105, 2024.

[21] A. Askarzadeh, "A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm," Computers & structures, vol. 169, pp. 1-12, 2016.