I. J. Computer Network and Information Security, **2025, 5, 15-30**
Published Online on October 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijcnis.2025.05.02

Novel Data Compression and Aggregation Approach in WSN Using Enhanced Walrus Optimisation

Krishan Kumar*
Department of Electronics and Communication Engineering, Bhagat Phool Singh Mahila, Vishwavidyalaya, Khanpur, Sonepat, Haryana, India
E-mail: krishan.bpsmv@gmail.com
ORCID iD: https://orcid.org/0009-0003-8039-0159
*Corresponding author

Priyanka Anand
Department of Electronics and Communication Engineering, Bhagat Phool Singh Mahila Vishwavidyalaya, Khanpur, Sonepat, Haryana, India
E-mail: anand_priyanka10@yahoo.co.in
ORCID iD: https://orcid.org/0009-0002-8728-735X

Rajini Mehra
ECE department, DCRUST, Murthal, Sonipat, Haryana, India
E-mail: rajni.ece@dcrustm.org
ORCID iD: https://orcid.org/0009-0003-4766-1375

Received: 06 May 2025; Revised: 15 July 2025; Accepted: 21 August 2025; Published: 08 October 2025

Abstract: Wireless Sensor Networks (WSNs) play a crucial role in applications such as remote monitoring, surveillance, and the Internet of Things (IoT). Addressing the challenge of energy consumption is paramount in WSN design due to the finite nature of energy resources. In cluster-based WSNs, cluster heads (CH) perform vital tasks like data collection, data aggregation, and exchange with the base station. Therefore, achieving efficient load balancing for CHs is crucial for maximizing network longevity. Previous studies have considered load balancing with optimal CH selection, but the issue of data redundancy is not addressed. Data redundancy in processing and transmitting information to analysis centres significantly depletes sensor resources like (energy, bandwidth and such). This paper proposes a novel energy-efficient data aggregation approach with data compression termed C-EWaOA that is (Compression based Enhanced Walrus Optimization with a cognitive factor). The non-negative matrix factorization (NMF) is used to compress the data and remove the redundant information. This way, the proposed data aggregation scheme reduces packet delivery Ratio, resulting in low data-rate communication. Simultaneously, data compression minimizes redundancy in aggregated data at CH, reducing resource consumption, leading to energy cost savings, and facilitating the transmission of a compact data stream in the communication bandwidth. The proposed approach shows a 0.606% improvement in network lifetime compared to the approach without compression and 68.01% of energy consumption. Notably, it achieves a reduction of 78.57% in packet loss ratio compared to the state-of-the-art FEEC-IIR model. Thus, the proposed approach shows remarkable improvement in energy-efficient data aggregation with data compression in WSN showcasing its prominence in IoT-based applications.

Index Terms: IoT-WSN, Data Aggregation, and Compression-based Enhanced Walrus Optimization with a Cognitive Factor.

1. Introduction

WSNs are integral to Internet of Things (IoT) applications, playing a crucial role in diverse fields such as health, Industry 4.0, and more [1]. Sensor Nodes (SNs) within WSNs are pivotal in the IoT's sensing layer, tasked with perceiving the physical environment and collecting data. The network comprises numerous tiny sensor nodes operating on limited battery life, emphasizing the need for effective energy management [2]. Clustering emerges as a widely used technique in WSNs to conserve energy [3,4].

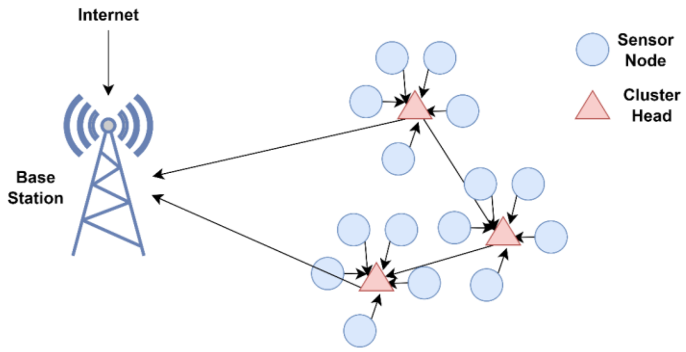

Fig.1. Cluster-based WSN (the arrows illustrate the communication)

In the cluster-based architecture of WSN as a scene in Fig. 1, the sensor nodes are grouped together as a single cluster where a cluster member is responsible for the communication denoted as the cluster head (CH) [3]. As seen in Fig. 1, The CHs are responsible for processing the local data generated in the WSN architecture from the members in the cluster known as sensor nodes and processing it directly to the sink [3]. This architecture reduces energy consumption such that only the CH is responsible for the data aggregation and routing process. This way, the bandwidth is conserved as the nodes in the cluster avoid sharing redundant information among them [4]. As there are advantages at some point when the data becomes high, especially in applications like IoT and Internet of Vehicles (IoV), the load on the CH to communicate becomes overbearing and causes rapid depletion of energy for SN, causing the node to die. Hence, an optimal way to overcome this issue is through load-balanced clustering and efficient CH selection. The two main issues addressed in the study are load-balanced clustering and energy-efficient CH selection to reduce the energy cost and improve the network performance. Such an issue is addressed using a multitude of optimisation and heuristics approaches in recent studies [5-7] as discussed in Section 1.1.

1.1. Previous Works

To overcome the load balancing issue in WSN to promote efficient data aggregation, the previous studies proposed clustering and optimal CH selection. A metaheuristics-based approach was proposed by [4] for clustering energy-efficient clusters. The parameters of residual energy of nodes, reduced energy consumption, distance between the nodes, and size of clusters were considered for decision-making of clustering and CH selection. However, here in the study by [4], the nodes are homogenous, and in real-time, the data processing for such a network is high. Researchers used the PSO algorithm to address the limited energy resources in CH in WSN [5]. The parameters distance between the CH and Base station, number of relay nodes and relay load factors were considered for clustering and CH selection. Researchers in [6] proposed using a Chaotic Genetic Algorithm (CGA) for generating the CH in homogenous nodes in WSN. Inter and intra-energy consumption of the network, and distance between the nodes were considered for CH selection in [6]. Work in [7] discussed the primary concern as the limited energy of battery-powered SN, which can lead to WSN network degradation or failure over time. To address this, it employed the Butterfly Optimization Algorithm (BOA) for optimal cluster head selection and the ACO for identifying the optimal route between the CH and the base station. To select the CH, [7] utilized the residual energy of SN, the distance between the nodes, the node degree, and the distance of nodes to the BS. The authors in [7], expressed that the energy consumption is directly proportional to the distance between nodes. To address the limited power of the SN in WSN, multiple attribute decision-making methods for CH selection is also suggested [8]. This approach primarily used residual energy and distance from nodes to CH and BS, and the clustering objective is related to reducing the energy consumption for load balancing. Similar to other studies, researchers in [9] utilized improved grey wolf optimisation (GWO) for CH selection. The authors emphasized the imbalance between other algorithms' exploration and exploitation phases. However, even these algorithms do not address the clustering load balancing issue in heterogeneous WSNs. The CH selection parameters in [9] are sink distance, residual energy, average intra-cluster distance, and CH balancing factor. The study in [10] examined two methods: the Hybrid Butterfly and Ant Colony Optimization algorithm with a static sink node and mobile sink node (HBACS, HBACM). The objective of these approaches is to enhance the selection of CH and provide energy-efficient routing in order to minimize energy usage and maximize the lifespan of the network. Similar to other studies, here as well the residual energy, distance, and node degree are considered for CH selection, whereas for cluster formation, SN are clustered with minimal distance, and maximum residual energy is clustered.

Table 1. Overview of the previous studies

Paper	Efficient Load balancing	Data redundancy through data reduction or compression techniques	Effective parameters for CH selection
[4]	✓	✕	Residual energy, Distance, Size of clusters
[5]	✓	✕	Distance to CH, Number of relay nodes, Relay load factors
[6]	✓	✕	Inter and intra energy consumption, Distance between nodes
[7]	✓	✕	Residual energy, Distance between nodes, Node degree, Distance to the BS
[8]	✓	✕	Residual energy, Distance from nodes to CH and BS
[9]	✓	✕	Sink distance, Residual energy, Average intra-cluster distance, Balancing factor of CH
[10]	✓	✕	Residual energy, Distance, Node degree
[11]	✓	✕	Distance between nodes, Gateways for cluster formation
[12]	✕	✓	NMF for data representation
[13]	✕	✓	NMF for outlier detection
[14]	✓	✓	NMF and 2-SVD-QR for data compression
[15]	✕	✓	NMF for data reduction
[16]	✓	✕	Residual energy of nodes
[17]	✓	✕	Nodes' initial energy, Remaining energy, Residual energy for CH selection
[18]	✓	✕	Residual energy of nodes
[19]	✕	✓	Energy cost, compression ratio
[20]	✓	✕	Residual energy of nodes, total energy
[21]	✓	✕	Intra cluster distance and average sink node
Proposed model	✓	✓	Based on minimizing the fitness function based on 1. Load balancing of CH, 2. Energy consumption between nodes (CM) and 3. energy consumption of CH after compression.

The literature review emphasizes the importance of clustering strategies in WSNs for minimizing energy consumption and enhancing network lifespan. Although current strategies address issues like energy efficiency and load balancing for data aggregation, there are research gaps for comprehensively tackling data redundancy or the energy cost of data storage and processing. Data compression techniques ([12-15], [19]) have been proposed to reduce data redundancy, but their integration into load balancing and CH selection strategies is still lacking. Most CH selection approaches prioritize residual energy ([4-7]), without considering other factors like energy required for data compression, distance, and trust, leading to suboptimal network performance. Thus, the key difference between the previous works and the proposed solutions is given in Table 2.

Table 2. Explanation of problems identified and proposed solution

Problems identified	Proposed solution
The current studies work towards load balancing in WSN but ignore the high amount of data aggregated at selected CH leading to high energy consumption [5,6] [9] [11] (as evident from Table 1).	The data aggregation is improved is incorporated with compression of data from SN which takes place at CH.
In WSNs, clustering strategies help minimize energy usage but fails to address the uneven load distribution across the network. Most of the studies like in [4-9] and such as seen in Table 1 address the load balancing issue but fails to address the data redundancy.	In the proposed model, a compression-based load balanced clustering scheme is employed. Different from the previous studies, in the proposed model, the overloading of data at CH is addressed by employing non-negative matrix factorization scheme.
For energy-efficient CH selection is carried out based on multiple parameters like residual energy [4,5] [16] [18], node degree [7] or energy consumption [6]. However, CH is selected based on heuristic algorithms like PSO [5], GWO [9], and hybrid models like in [10]. These models do not guarantee for optimal global solution and are based on random search schemes; this gives sub-optimal results for CH selection	Different from the previous models where CH selection is based on single metaheuristics which are limited to the local search space and hybrid models which are not domain specific. In this study, a novel enhanced nature-inspired metaheuristics algorithm is proposed. EWaOA balance both exploration and exploration stages with the addition of cognitive factor in the exploration phase, improving the search space globally for optimal CH selection.

1.2. Motivation and Contributions of the Study

In this work, the studies are reviewed for two main processes of data aggregation: Load balancing-based clustering and CH selection. Some of the commonly used optimisation algorithms for CH selection to address load balancing are the Genetic algorithm (GA) [6], PSO (Particle Swarm Optimisation) [5], and ACO (Ant-colony Optimisation) [7]. Due to the stochastic and random search nature of metaheuristic algorithms like PSO, GA, and DE, the solutions derived from these algorithms are referred to as quasi-optimal solutions [25]. Such algorithms initiate random feasible solutions in the search space, then proceed with the iterative process to improve the initial solutions. Such issues are addressed using hybrid optimisations like in [11] GWO-GA (Grey wolf and Genetic algorithms), [16] using enhanced penguin optimisation and in [21] Quasi oppositional based Jaya algorithm and [22]. Thus, there is a need to find the balance between exploration and exploitation phase is important to develop optimal decisions which are feasible for all clustering scenarios (dense and weak) in WSN [26]. Another task is energy-efficient CH selection, from the literature residual energy is widely used along with sensor node related information and distance metrics. Thus, the main limitations to be addressed are,

- In WSNs, clustering strategies help minimize energy usage, but fails to address the uneven load distribution across the network. Most of the studies like in [4-9] and such as seen in Table 1 address the load balancing issue but fails to address the data redundancy as explained in Section 2.1.
- For energy efficient CH selection is carried out based on multiple parameters like residual energy [4,5] [16] [18], node degree [7] or energy consumption [6]. However, CH is selected based on heuristic algorithms like PSO [5], GWO [9], and hybrid models like in [10]. These models do not guarantee for optimal global solution and are based on random search schemes, which give sub-optimal results for CH selection.

Motivated by the concept of NMF for redundancy and to address the issues above, the paper proposes 1. The problem of CH selection is addressed through Novel Compression-based Enhanced Walrus Optimisation (C-EWaoA) that incorporates data compression in diverse problem spaces. 2. A multi-objective-based CH selection. Thus, the two main contributions are given as,

- To propose a data redundancy aware data aggregation in WSN. The objective of uniform node distribution in WSN and data redundancy at CH is achieved through Non-Negative Matrix Factorization for load-balanced clustering. This is done by implementing lossless matrix factorization technique to compress data at the CH, thereby lowering CH energy consumption and increasing the system lifespan.
- To propose a novel Compression-based Enhanced Walrus Optimisation (C-EWaoA) model for optimal CH selection, where a cognitive factor is applied in the exploration phase of WaoA for improving the global search space. The CH selection is based on a multi-objective fitness function based on the energy consumption of nodes, CH and energy consumption after compression. The goal is to enhance the lifespan of active nodes and reduce energy consumption in both directions of communication between nodes and the CH.

The remaining paper is divided into Sections, where in Section 2 the proposed system model and problem statement are discussed. In Section 3 the proposed approach is detailed. In Section 4 the experiments and discussions are made on the results obtained.

2. Proposed Model

In this Section, the basic description of the proposed structure of the framework is provided and the problem formulation is explained. The uneven distribution of loads among the Sensor nodes is addressed by proposing efficient load balancing focusing on reducing the energy costs and improving the network lifetime. The proposed novel data aggregation scheme here is divided into two stages: Clustering and CH selection.

2.1. Problem Formulation

Stage 1: Clustering for load balancing

The main challenges for data aggregation in WSN are achieving energy efficiency, load balancing, and reducing data redundancy [7,12]. Sensor nodes in WSN often have limited resources, particularly when it comes to battery life. Data transmission between these nodes incurs a substantial energy cost, and frequent communication might result in the exhaustion of batteries. In the cluster-based IoT WSN architecture as illustrated in Fig. 1, it is necessary to balance the load among the nodes to avoid excessive use of energy by a single CH in clusters. Such overloading at CH results in packet delay reducing the lifetime of the network. Hence, in this study, the load-balancing problem is addressed by clustering the SN in the IoT-WSN area. Let N be randomly distributed sensor nodes in the WSN area and the data generated from the Sensor node is represented as $Y_{j\in(1,2.....N)} = (y_1, y_2 \dots y_N)$ and mathematically given as

$$Y = Y_{N\times N} = \begin{bmatrix} y_{1,1} & \cdots & \cdots & y_{1,k} & y_{1,o-1} & y_{1,o} \\ y_{2,1} & \cdots & \cdots & y_{2,k} & \cdots & y_{2,o} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{O-1,1} & \cdots & \cdots & y_{O-1,k} & \cdots & y_{O-1,o} \\ y_{O,1} & \cdots & \cdots & y_{O,k} & y_{O,o-1} & y_{O,o} \end{bmatrix} = WH = W_{Nk}H_{kN} \tag{1}$$

From equation 1, it can be seen that the data is explored a factorization, where the main objective of NMF employed is to remove the redundant information generated by decomposing the matrix Y in two matrices W_{Nk} and H_{kN}, where $N < k$ (k are the clusters formed). The fine-tuned version of the matrix H_{kN} where the redundant information is eliminated, is applied for k-means clustering and this way k clusters are generated. The realization of NMF is briefly discussed in [14] and adopted here.

Stage 2: Novel Enhanced Optimisation algorithm-based CH selection: C-EWaOA

The IoT-WSN area consists of a Base Station (BS) with a stable energy source and Sensor Nodes (SNs) ($N = \{1,2,3 \dots n\}$, where n is the number of Sensor nodes in the area). Each SN is identified with a unique ID and is distributed randomly. The purpose of assigning SN as CH is to aggregate the data and transfer to the BS. At the base station the data is analysed processed and transmitted to CH, this data is received by the Cluster members. Hence, the data aggregation takes place at CH. In such cases, where the data aggregation is huge like in IoT applications, the battery life of CH node depletes faster causing it to die before it can transmit the data resulting in poor network performance. Although the proposed load balancing ensures efficient clustering of SN, maintaining the durability of the CH is to be addressed. In this paper, the second stage is proposed to select an optimal CH which reduces the energy consumption and transmission delay. This problem of CH selection is formulated as an optimisation problem.

The proposed framework introduces optimisation (EWaoA) that incorporates the data compression technique Non-matrix Factorization (NMF) to determine the optimal CH, this is illustrated in Fig. 2. The primary objective function f_1 is designed for load balancing in CH selection and is achieved by ensuring a uniform distribution of nodes across each cluster with the above process. The second objective f_2 is to minimize energy consumption for communication from CM to CH, by computing the data compression at the CH using the NMF method. Finally, the third objective f_3 is to minimize the energy consumed, when the compressed data is received by Sink from CH. Here the data is re-constructed at CH after receiving from the sink, to be transmitted to CM.

The three objectives mentioned serve as inputs to the proposed Compression EWaOA method, which aims to determine the optimal dynamic CH with a balanced load, as depicted in Fig. 2. The objective function relies on multiple fitness functions, and these objectives are connected to a shared sensor network attribute, namely distance. Consequently, they do not pose a Pareto front optimization challenge. To address this, the objectives are transformed into a weighted single objective problem within the proposed objective function for C-EWaoA, denoted as $f(x)$, and is expressed as follows,

Top of Form

$$f(x) = \sum_{i=1}^{3} C_i f_i = C_1 \times f_1 + C_2 \times f_2 + C_3 \times f_3 \tag{2}$$

Where C_i is the weight assigned to the fitness functions, the first fitness f_1 is given as,

$$f_1 = \left(1 - \frac{\mu_{load}}{CH_{max}}\right) + \left(\frac{CH\ loaded}{Total\ CH}\right) \tag{3}$$

This fitness function (f_1) in Equation 2, is the metric to evaluate the balanced load for the selected CH. It calculates the ratio of mean load μ_{load} of total CH from [24] and CH with maximum load (CH_{max}), which is given as, $\mu_{load} = \frac{\sum_{i=1}^{m} Load\ (CH_q)}{cluster\ heads}$. Whereas, $Load\ (CH_q) = N \times \frac{E_{remain}\ (CH_q)}{E_{initial}\ (CH_q)}$ is based on the number of SN in the System model (N), the initial and remaining energy of CH that is $\frac{E_{remain}\ (CH_q)}{E_{initial}\ (CH_q)}$. The primary goal of f_2 is to measure the energy utilized by individual nodes while transmitting data. The energy usage is affected by the proximity of the sender and receiver nodes d_0. It is given as,

$$f_2 = E_T(N, d) = \begin{Bmatrix} N \times E_{elec} + N \times \in_{fs} \times d^2, & d < d_0 \\ N \times E_{elec} + N \times \in_{mp} \times d^4, & d \geq d_0 \end{Bmatrix} \tag{4}$$

This fitness function provides a quantifiable measure of the energy consumed by nodes during data transmission, considering both free space and multipath energy characteristics ($\in_{fs}, \in_{mp}$) based on d_0. Similarly, to f_2, after compression the proposed f_3 computes the energy consumed by individual CH. It is given as,

$$f_3 = E_T(CH, d) = \begin{Bmatrix} CH \times E_{elec} + CH \times \in_{fs} \times d^2, & d < d_0 \\ CH \times E_{elec} + CH \times \in_{mp} \times d^4, & d \geq d_0 \end{Bmatrix} \tag{5}$$

Based on the objective function given in Equation 1, the novel C-EWaOA is optimised for CH selection.

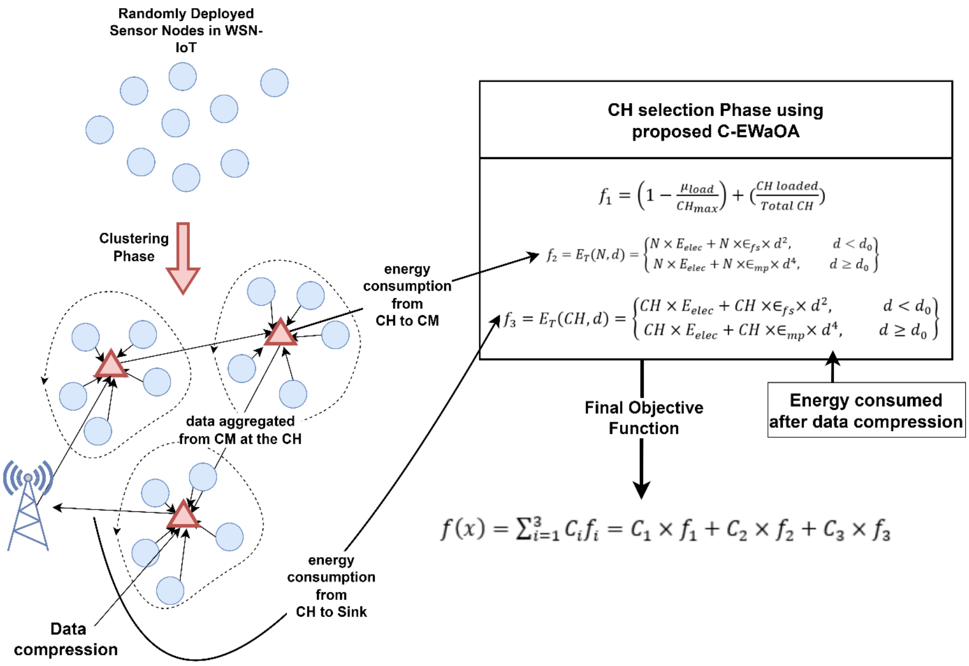

Fig.2. Illustrationof flow of the structure of proposed framework

2.2. *Proposed Enhanced Optimisation Algorithm for Optimal CH Selection*

Based on the problem formulation, the objective function formulated in Equation 1 is realized for optimal CH selection as a minimum $f(x)$. This is achieved with the training of the proposed C-EWaOA. The iterative process in the proposed novel C-EWaOA is responsible for the selection of optimal CH based on the new position sets of walruses for each cluster k as illustrated in Fig 3. The decision variables m is represented as a number of clusters k such that, $m = k$. For every iteration in T, the objectives function formulated in Equations 2, 3 and 4 are calculated and stored as candidate solution matrix X as explained in Equation 5 of size $N \times m$, the solutions are stored in a matrix X and with Equation 11 the best CH is elected and considered as the best solution so far in C-EWaOA.

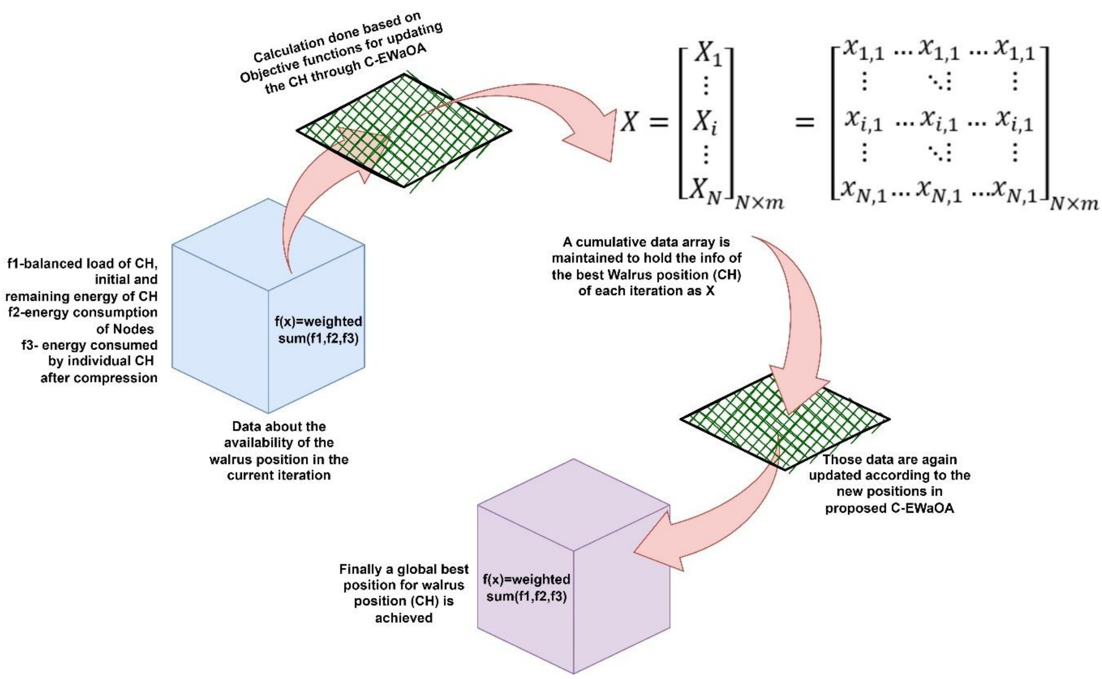

Fig.3. The generic process of CH selection in EWaOA

Description of Enhanced WaOA

Most of the studies introduce metaheuristic algorithms like PSO, GA, and DE, but due to the stochastic and random search nature of metaheuristic algorithms, the solutions derived from these algorithms are referred to as quasi-optimal solutions [25]. Such algorithms initiate random feasible solutions in the search space and then proceed with the iterative process to improve the initial solutions. In such issues, balancing the exploration and exploitation phase is important to develop optimal decisions [26]. To overcome this issue, the paper in [27] proposes a new optimisation algorithm inspired by Walrus, named WaOA (Walrus Optimisation Algorithm). The optimisation is inspired by the nature of "Walrus" mammas around the North Pole. The social and natural life of the animal will be the inspiration of the algorithm [27]. The intelligent behaviours exhibited by walruses, which are mirrored in the Conventional Walrus Optimization Algorithm (WaOA) elaborated in [27], can be summarized into three key aspects: Co-ordination in feeding and Migration to Rocky Beaches, where the global search is carried out, then the predatory Response where the walruses make small dynamic changes showing the local search. The phases in the algorithm are depicted in Fig. 4 and the mathematical formulation of the conventional WaOA is given as,

The initial population is randomly initialized and is given by X and these are the positions of Walruses [27].

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N\times m} = \begin{bmatrix} x_{1,1} \dots x_{1,1} \dots x_{1,1} \\ \vdots \quad \ddots \vdots \quad \vdots \\ x_{i,1} \dots x_{i,1} \dots x_{i,1} \\ \vdots \quad \ddots \vdots \quad \vdots \\ x_{N,1} \dots x_{N,1} \dots x_{N,1} \end{bmatrix}_{N\times m} \tag{6}$$

Where N is the number of walruses and m are the decision variables. The decision values of the objective functions are stored in F and given as,

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N\times m} = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_i) \\ \vdots \\ F(X_N) \end{bmatrix}_{N\times 1} \tag{7}$$

The conventional WaOA is designed to balance exploration and exploitation through its three distinct phases, each inspired by the natural behaviours of walruses. The feeding strategy, inspired by the walrus's search for food, explains the exploration. The walrus with the best objective function value is considered the strongest, guiding others to explore different areas of the search space, and promoting global search. It is explained as,

$$X_{P1_{i,j}} = X_{i,j} + rand_{i,j} \cdot (SW_j - I_{i,j} \cdot X_{i,j}) \tag{8}$$

$$X_i = \begin{cases} X_{P1}, & if\ F_{P1_i} < F_i \\ & X_i\ else \end{cases} \tag{9}$$

$X_{P1_{i,j}}$ is the new position for i^{th} Walrus in j^{th} iteration during this stage and SW represents the strongest walrus position. Whereas in the exploration phase, the migration phase mimics the walruses' movement to new areas, enhancing exploration by encouraging walruses to explore the positions of other walruses in the search space.

$$X_i = \begin{cases} X_{P2_{i,j}} = X_{i,j} + rand_{i,j} \cdot \left(X_{k,j} - I_{i,j} \cdot X_{i,j}\right),\ F_k < F_i & if\ F_{P2_i} < F_i and\ k \neq i \\ X_i & else \end{cases} \tag{10}$$

Where, $X_{k,j}$ and F_k represents the position and Objective function value of a randomly selected walrus (k).

In the Exploitation phase, local search is carried out around the position of the walrus simulating its behaviour of escaping and fighting the predators. It is given as,

$$X_i = \begin{cases} X_{P3_{i,j}} = X_{i,j} + (lb^t_{local}, j + (ub^t_{local}, j - rand \cdot lb^t_{local}, j)),\ if\ F_{P2_i} < F_i \\ X_i\ else \end{cases} \tag{11}$$

Whereas, the local lower and upper bounds for j^{th} variable are given by lb^t_{local}, j, ub^t_{local}, j.

However, as discussed in Section 1.2, for any metaheuristic algorithm to be truly effective, it must explore the solution space thoroughly to identify a globally optimal solution, ensuring it adapts to the varying characteristics of nodes in a WSN without getting trapped in local optima. Thus, to enhance the efficiency of algorithmic hybridization, the Social Cognitive Theory is applied here, as introduced by [28]. This theory, widely applied in education, communication, and psychology, posits that an individual's knowledge acquisition is directly linked to their experiences, interactions, and external influences. Consequently, individual actions are guided by attitudes, learning from the trial and error of others, predicting outcomes based on experiences, and efficiently attaining their goals.

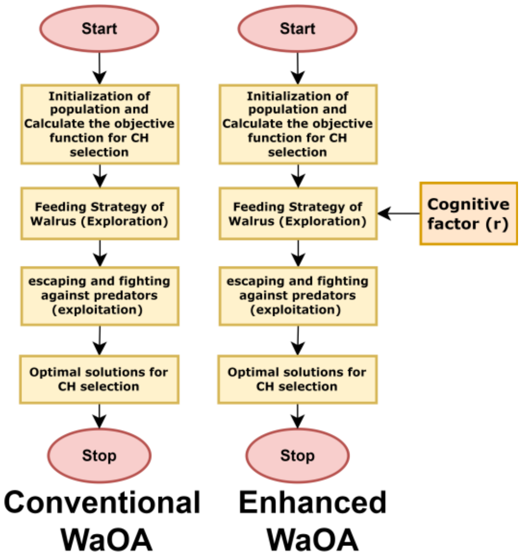

Fig.4. Illustration of working of Conventional WaOA and proposed enhanced WaOA

In light of this, the cognitive component employs the deduction of the local best position from the randomly chosen current set of CH positions.

$$X_{P2_{i,j}}: x_{P2_{i,j}} = \begin{cases} X_{i,j} + rand_{i,j} \cdot \left(X_{k,j} - I_{i,j} \cdot x_{i,j}\right) + (\mathbf{1} - \boldsymbol{r})(\boldsymbol{local}(\boldsymbol{j}) - \boldsymbol{I_{i,j}}.\boldsymbol{x_{i,j}}), & F_k < F_i \\ X_{i,j} + rand_{i,j} \cdot \left(X_{k,j} - I_{i,j} \cdot x_{i,j}\right) else & \end{cases} \tag{12}$$

Thus, the exploration stage given in Equation 7 is improved with third component as cognitive factor to increase the search space as depicted in Fig. 4. Hence, the equation 7 is Enhanced as equation 11 for CH selection.

2.3. Enhanced Optimal CH Selection

The proposed approach for CH selection is Enhanced with cognitive factors as given in Equation 4. The cognitive factor r [28], helps the exploration phase of WaOA to improve its ability to optimize better for non-linear functions. This drives the WaOA to converge towards the optimal solution by uploading the heterogeneity among its constituents and prevents it to fall in local minima.

The estimated values of the objective function for suggested candidate solution for CH selection problem are formulated in Equation 1, and is given by matrix with these solutions as F [27]. In the proposed model the objective function for the election of CH is given by vector F from Equation 6, is formulated as,

$$F = F(X_{i\in 1,2..N}) = min\ (F_i^{P_i}) \tag{13}$$

That is if m is the number of clusters (m=k), then through the iterative process of EWaOA the objective functions in Equations 2, 3 and 4 are calculated and stored as X in Equation 5 with $N \times m$ possible candidate solutions (walrus) and stored as F and through equation 11 the CH is selected.

Algorithm 1. Pseudocode for Optimal CH selection with Novel EWaOA

Input: Number of Walrus N, Iterations T
1. Load the sensor data.
2. Apply K-means clustering and generate the clusters k
3. Initialize the random positions of walrus X as in Equation 5 (i.e., CH in our proposed model)
4. **For** $t = 1{:}T$

a. Update the CH based on the objective function calculated in Equation 1

b. **For** $i = 1{:}N$

i. Calculate the new j^{th} CH from X using equation 11 where the exploration phase is Enhanced with cognitive factor. # **Enhanced Exploration Phase in** EWaOA

ii. Update the position of i^{th} CH in F in Equation 8.

iii. Choose the immigration destination for the $i^{th}$and calculate the location of CH using equation 9 **#migration phase in** EWaOA

iv. Update the position in F based on the conditions from Equation 99

v. Calculate the new position for CH through Equation 10 and update it in F **#Exploitation Phase in** EWaOA

c. End **For**

d. Save the best solutions so far as in F
5. End **For**
6. Select the Best CH from the matrix of solutions as formulated in Equation 6.
End of algorithm
Output: Optimal CH

As per Enhanced equation 11, the current set of CHs is selected by EWaOA with the incorporation of terms $(\boldsymbol{local(j) - I_{i.j}.x_{i,j}})$, where the $local(j)$, the local position is selected randomly from X_i set of candidate solutions of CH for each cluster m over the j^{th} iteration. Hence, this process is termed as Optimal CH selection through IWaOA. Hence, the updated set of values in Equation 6 so far after the iterative process of proposed IWaOA explained in Algorithm 1 are considered as the optimal set of CHs. With the proposed EWaOA the model convergence increases but due to higher payload the energy of CH depletes faster and it dies. This problem is formulated by introducing data compression at CH as explained in Section 2.1. the compression of raw data at CH with NMF and re-construction at CH after receiving is achieved with promising MSE (mean square error). The selection of an appropriate compression ratio relies on the Mean Squared Error (MSE) between the decompressed and reconstructed data [19]. Higher compression ratios may lead to a significant reduction in raw data size but come at the expense of increased data loss during the reconstruction process.

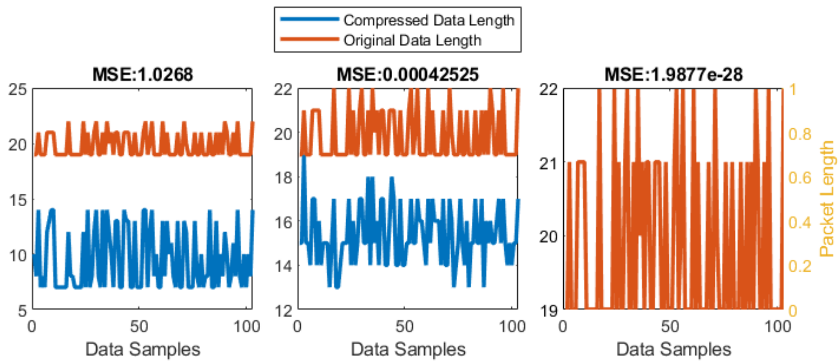

Fig.5. Illustration of original data at CH and compressed data with MSE analysis using proposed NMF data compression

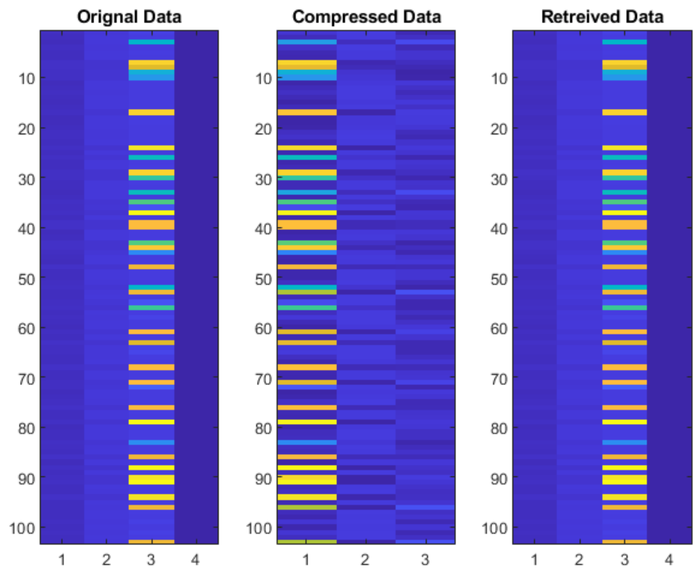

Fig.6. Visualization of the original data at the sink, compressed data intended for transmission, and the reconstructed data upon reception at the sink

Fig. 5 displays 3 different cases corresponding to 3 distinct low ranks in the NMF factorization. The first case involves compression with a rank of 2, resulting in a high reconstruction MSE of 1.0268. On the other hand, the third case achieves a very low MSE for rank 4, with the length of compressed data negligible compared to the original data length. The second case, with an optimal MSE of 0.0004, is chosen for transmission, striking a balance between the compression ratio and reconstruction error, as depicted in Fig. 6. The MSE of 0.0004 is identified as a trade-off between compression ratio and reconstruction error. In the first case, with an MSE of 1.026, data loss during reconstruction is higher, while in the third case, reconstruction is nearly impossible due to the excessively high data compression ratio. This observed MSE of 0.0004 is considered as an optimal compromise for the proposed transmission, this trade-off is demonstrated in Fig 6.

3. Experiments and Discussion

3.1. Simulation and Evaluation Parameters

The simulation of proposed model is simulated using MATLAB and the. The proposed model is evaluated based on network lifetime, packet delivery ratio and residual energy parameters. For Simulation, the proposed system model is initialized with 500 SNs showcasing the IoT environment with a Random topology pattern with an omnidirectional antenna. Each node starts with an initial energy of $0.5\,J$, and the transmission distance is set at 20 m. The packet size utilized in the simulation is 64 bytes, and the data rate is established at 14 packets per second. With the implementation of the clustering process on the SN, it is illustrated in Fig. 7 that 6 clusters are generated.

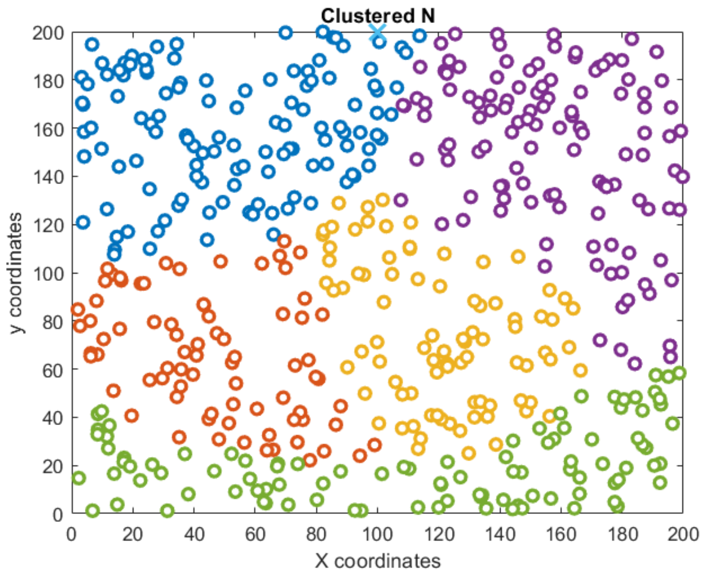

Fig.7. Clustering of SN: the different colours (Red, green, blue, yellow and purple) indicate each cluster

The suggested model is tested across the metrics, on the basis of network parameters: packet delivery ratio (PDR), network lifetime [4], and residual energy. The analysis of the results demonstrated that it has improved the network lifetime, and the idea of data compression through NMF is implemented. There are four test cases that compare each network metric. The test cases are indicated by the plot legends, which stand for 1) the proposed model that is EWaOA with data compression (C- EWaOA), 2) conventional WaOA, 3) Conventional WaOA with data compression (C-WaOA), and 4) Enhanced WaOA (EWaOA).

3.2. Analysis of Results

This research contributes a novel methodology for optimizing load balancing and energy consumption in Wireless Sensor Networks (WSNs) by integrating data compression with CH selection. The innovative use of the cognitive-enhanced Walrus Optimization Algorithm (WaOA) advances the theoretical understanding of optimization processes in WSNs, providing a robust framework that can be adapted for further research in related fields such as distributed computing and IoT systems.

In this paper, the node distribution is random and the count of nodes is 500 indicating the dense nature of WSN-like IoT applications. As discussed in Section 2.1, NMF-based clustering is performed on the sensor nodes and five clusters are produced to test the uniformity of the number of nodes (to show the load-balanced clustering scheme proposed by NMF) which is 97, 96, 99, 101 and 107 in each cluster distributed in IoT scenario respectively. The lifetime of nodes over the varying test cases is simulated and given in Fig. 8. Live nodes is an important metric for measuring the robustness and efficiency of the WSN network. When a node's energy supply becomes low, it may lose the ability to communicate or conduct sensing operations. The suggested compression-based EWaOA outperforms other state-of-the-art models in terms of node lifetime, as shown in Fig. 8.

Alive nodes are an important metric for measuring the robustness and efficiency of the network. The robustness of the WSN system with IoT application is dependent on maintaining a sufficient number of “alive” nodes (functioning sensors) to ensure accurate data collection and timely responses. The experiments in Fig. 8, show that C- EWaOA shows a 1.02% improvement in network lifetime compared to EWaOA, showcasing that the proposed model strategically avoids selecting cluster heads with lower residual energy, thereby enhancing the network's longevity. When a node's energy supply becomes low, it may lose the ability to communicate or conduct sensing operations for example, in applications like healthcare, Prolonged node lifetime ensures continuous health monitoring without frequent battery replacements. With the application of a compression mechanism for load balancing, the proposed model is able to save energy by removing redundant information. Which in turn leads to less frequent exhaustion of batteries in sensor nodes suggesting that the compression-based EWaOA outperforms EWaoA without compression in terms of node lifetime, as shown in Fig. 8. Least alive nodes are captured by the conventional WaOA compared to enhanced WaOA with cognitive factor. This is because the factor helps maximize the search space as seen in Equation 11.

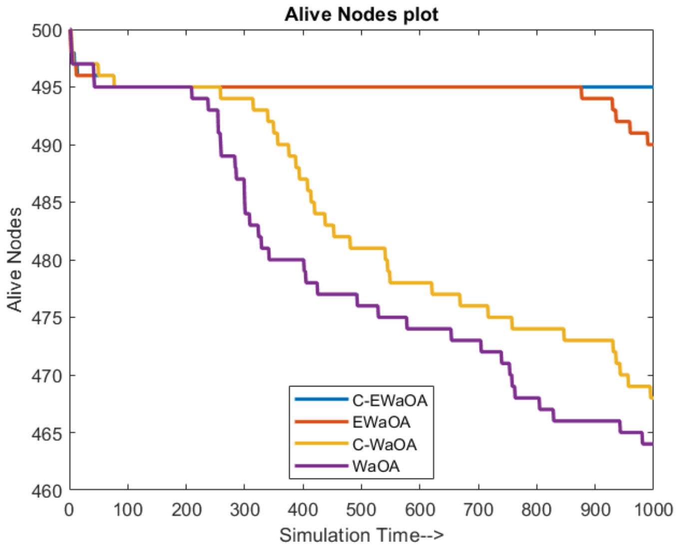

Fig.8. Graph illustration of number of nodes alive over the simulation time regarding the comparative network lifetime analysis

To achieve efficient load balancing, it is also important that the nodes (CH and CM) preserve the energy after the data is transmitted to improve the network performance. With a high number of alive nodes, the proposed model (C-EWaOA) is able to prevents the selection of CHs with low residual energy, which increases the network's lifetime. In every case that was studied in Fig. 9, the network lifetime is consistently improved when comparing compression-based and non-compression models.

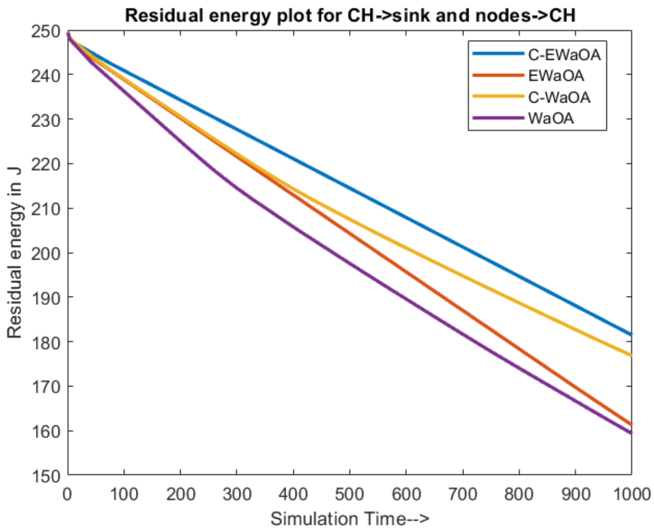

Fig.9. Comparative Analysis for the impact of Residual energy in Joules over the simulation time

In applications like IoT, where end-users rely on the data collected or transmitted by the WSN, a high network lifetime ensures a consistent and reliable service. To achieve this the proposed model applied compression on the data aggregated at CH, this leads to a 12.53% increase in residual energy compared to EWaOA without compression and nodes can continue functioning for longer periods, maintaining network reliability. This is because there is a lot of energy wasted. After all, processing uncompressed data requires a lot of power from the nodes. Efficient use of energy is also crucial in battery-powered networks like WSN. Such C-EWaOA, employing the suggested objective function, demonstrates a 2.61% improvement compared to C-WaOA. This is because the cognitive factor prevents the CH selection process from prematurely converging to suboptimal solutions by promoting diverse exploration. Furthermore, the results extend to Fig. 10, showcasing a consistent positive trend in the volume of successfully delivered packets. This graphical representation illustrates the effectiveness of the proposed model, C- EWaOA, in delivering a higher quantity of packets to the base station when compared to alternative optimization approaches.

The proposed model outperforms conventional methods. Specifically, C- EWaOA demonstrates a notable improvement of 0.08% and 2.34% in packet delivery compared to C-WaOA and Enhanced WaOA (EWaOA), respectively, according to the specified objective. It is noteworthy that even in the absence of data compression, EWaOA exhibits enhanced packet delivery owing to its superior convergence capabilities during the exploration phase. The introduction of data compression to C-EWaOA leads to an additional improvement of 0.08% over C-WaOA without compression, attributing this enhancement to Non-Negative Matrix Factorization's (NMF) favourable compression ratio and its role in reducing the reconstruction error, thereby contributing to the Enhanced packet delivery observed in the proposed scheme. In real-time applications, this means that a higher volume of critical data can be successfully transmitted to the base station, reducing data loss and ensuring more accurate and timely decision-making. The introduction of data compression further amplifies these benefits, making the C-EWaOA model particularly well-suited for deployment in battery-powered, resource-limited networks where energy efficiency and data fidelity are paramount.

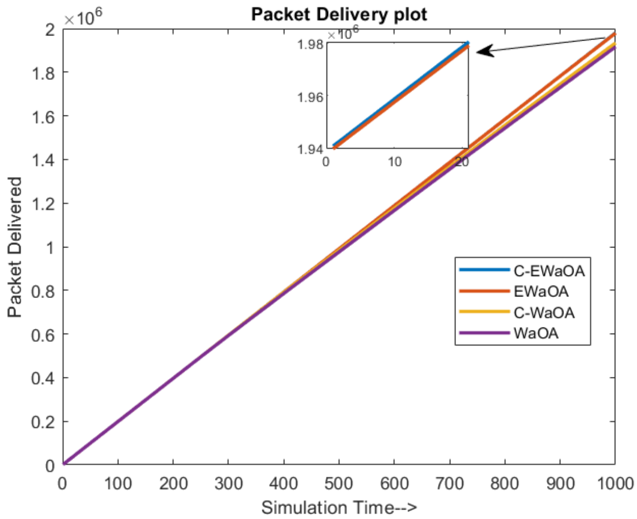

Fig.10. Comparative analysis of PDR over the simulation time

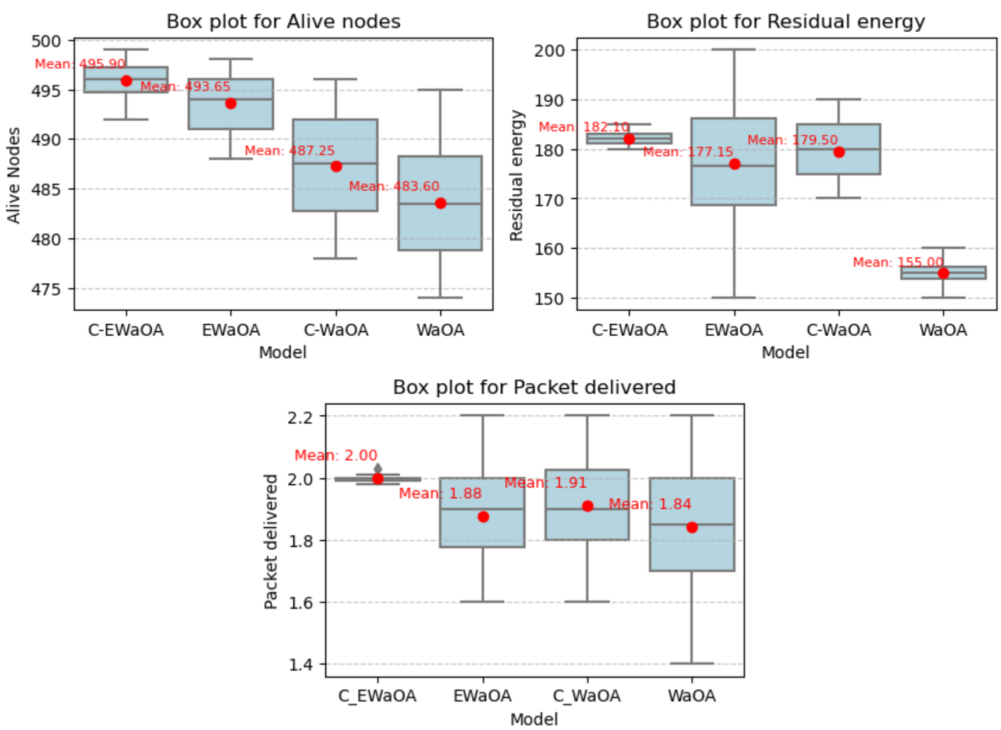

Fig.11. Statistical analysis

Statistical Analysis of the Proposed Model

In this section, the test cases are run and simulated for 20 trials with the same set of parameters for testing each variant and statistical results are given in Fig. 11. The box plot for C-EWaOA has the smallest interquartile range (IQR) and whiskers, indicating the least variance compared to other models. It is depicted from the box plot graph for Alive nodes, the proposed model shows a low standard deviation of 1.84, which is approximately 62%, 196% and 223% less compared to EWaOA, C-EWaOA and WaOA. The best value of Alive nodes recorded for the proposed model is 499 and the worst value is 492 along the simulation with no outliers detected.

The proposed model shows highest number of alive nodes in the WSN network. The mean value of the proposed model (C-EWaOA) in 20 trails is 495.90 which is 0.45%, 1.78% and 2.54% improved compared to EWaOA, C-WaOA and WaOA respectively.

This indicates that the data compression and the enhancements in the proposed model (C-EWaOA) significantly improve the number of alive nodes, suggesting better network longevity. In the case of residual energy, the proposed model shows a low standard deviation of 1.3 and the model shows 185J as the best value along the simulation and records 180J as the worst value. As seen in Fig. 11, the proposed model shows an improvement of approximately 2.79%, 1.45% and 17.48% compared to EWaOA, C-WaOA and WaOA in the case of residual energy (mean values). In remote areas, where replacing batteries is difficult, prolonging the network's lifespan is crucial. The residual energy of C-EWaOA makes it ideal for such applications as it shows a 947%, 369.23% and 172.3% decrease in standard deviation compared to the other three variants.

Whereas, in the case of the packet delivered C-EWaOA shows 0.0124 of standard deviation with the best value recorded at 2.03 and the worst at 1.98. It shows improvement of approximately 6.38% over EWaOA, 4.71% over C-WaOA and 8.7% over WaOA. The box plot is narrow as seen in Fig. 11, indicating a lower standard deviation, which is 1222.58%, 1279% and 1658% less compared to other three variants, suggesting that the model is able to deliver a consistent number of packets and shows consistent results over the simulation.

Whereas in the case of EWaOA, suggests more variability in the number of alive nodes, possibly due to the lack of data compression, leading to less efficient energy consumption. In the case of C-WaOA, WaOA has a high variance (29.69, 35.34), indicating fluctuations in the alive nodes throughout the simulation making it highly deviated from the mean value (actual value). In the residual energy box plot analysis, the standard deviation is low for the proposed model and WaOA, however, the mean is lower in the case of WaOA showing that this variant is not able to maintain a prolonged network lifespan throughout the simulation. This suggests significant variability in network performance recorded at 185, 37.25 and 12.50 for EWaOA, C-WaOA and WaOA respectively compared to the proposed model which is 1.69.

Similarly, in the case of packet delivered, the proposed model shows less variance of values throughout the simulation given as 0.000154. whereas, for the three variants (EWaOA, C-WaOA and WaOA) there is the high variance of packet delivered value for simulations which is 0.0269, 0.0289 and 0.0474 respectively.

The ability of the proposed model to maintain a higher number of alive nodes ensures the network can cover a larger area for a longer time. In real-time applications, such as environmental monitoring, healthcare, or military surveillance, where data aggregation in WSN is critical, the C-EWaOA model's superior performance makes it highly suitable as it shows consistent analysis over 20 trials.

3.3. Comparative Analysis with State-of-Art Models

In this Section, Table 3 presents a comparative analysis of the proposed model for 500 nodes, against state-of-the-art models in energy consumption, Packet Delivery Ratio (PDR), and Packet Loss Ratio. The considered algorithms include TCBDGA [31], HEED [32], and FEEC-IIR [33].

In [32], the proposed protocol introduces a probabilistic approach for CH selection, considering both energy levels and additional factors for advanced nodes. The authors in [31], propose a Tree-Cluster Based Data Gathering Algorithm (TCBDGA) designed specifically for WSNs with a mobile sink. The weight calculation considers factors such as average residual energy, distance to the Base Station (BS), and local node densities. The proposed model in this paper named the FEEC-IIR (Adaptive Fuzzy Rule-Based Energy-Efficient Clustering and Immune-Inspired Routing) protocol [33], addresses the significant issue of energy efficiency in wireless sensor networks (WSNs), particularly in the context of Internet of Things (IoT) systems. Energy status, Quality of Service (QoS) impact, and node location are identified as key criteria influencing CH selection. Compared to models [31-33], the proposed model achieves significant results with the novel addition of data compression for optimal CH selection achieving high PDR and less packet loss ratio

Table 3. Comparative analysis of the proposed model with state-of-art models in terms of energy consumption, PDR and packet loss ratio

Algorithms	Energy Consumption ($Joules$)	PDR (%)	Packet Loss Ratio (%)
TCBDGA [31]	145	88	8
C-EWaOA	***68.01***	***98.67***	***2.8***
HEED [32]	180	85	17
FEEC-IIR [33]	65	96	5

Because of compression at CH, the aggregation can be achieved with less consumed energy 68.01, the suggested model outperforms the other algorithms tested with 500 nodes in terms of PDR, outperforming TCBDGA by 10.81%, HEED by 13.85%, and FEEC-IIR by 2.7%. Whereas, the proposed model revealed a 53% and 62% decrease in energy use as compared to HEED and [31], respectively. Thus, the proposed model cannot only extend the lifespan of sensor nodes but also reduce operational costs, which is critical for sustainable and cost-effective IoT infrastructure. The model's ability to achieve a PDR improvement of up to 13.85% over existing algorithms ensures more reliable data transmission, which is crucial for applications such as environmental monitoring, smart cities, and healthcare systems where data accuracy and timely delivery are paramount.

4. Conclusions

In this research, C-EWaOA, a novel approach designed to tackle load-balanced data aggregation challenges in WSNs, with a particular focus on energy efficiency through compression within the Internet of Things (IoT) landscape is introduced. The core objective of optimal load balancing is achieved through a fitness function defined for CH selection, integrated with data compression techniques. The application of NMF during CH selection effectively reduces overloading and energy costs, leading to improved overall network performance. The proposed optimization process incorporates an enhanced version of the WaOA, augmented with a cognitive factor that strengthens the exploration phase. This enhancement helps prevent convergence to local minima and steers the algorithm toward more optimal solutions. Extensive simulations conducted in MATLAB validate the effectiveness of the proposed model, demonstrating significant improvements in network lifetime by achieving high alive nodes (1.02% higher compared to the model without compression), residual energy, and improved PDR (0.08% compared to without compression) and also compared to conventional WaOA, Conventional WaOA with data compression (C-WaOA), and Enhanced WaOA (EWaOA). Compared to existing state-of-art models such as TCBDGA, HEED, and FEEC-IIR proposed model. that is achieving 2.8% improvement in packet loss compared to TCBDA and 68J of low energy consumption, which is achieved at 500 nodes in the IoT-WSN environment. By achieving low packet loss, the proposed model ensures reliable aggregation maintaining the extended battery life of sensor nodes. While the results are promising, we recognize the importance of further exploration, particularly in the context of hybrid optimization models and scenarios involving varying node mobilities, which will be the focus of future research.

References

[1] Majid, M., Habib, S., Javed, A. R., Rizwan, M., Srivastava, G., Gadekallu, T. R., & Lin, J. C. W. (2022). Applications of wireless sensor networks and internet of things frameworks in the industry revolution 4.0: A systematic literature review. *Sensors*, *22*(6), 2087.

[2] Kumar, A., Kumar, V., & Chand, N. (2012). Energy efficient clustering and cluster head rotation scheme for wireless sensor networks. *International Journal of Advanced Computer Science and Applications*, *3*(5).

[3] Yarinezhad, R., & Hashemi, S. N. (2019). Solving the load-balanced clustering and routing problems in WSNs with an fpt-approximation algorithm and a grid structure. *Pervasive and Mobile Computing*, *58*, 101033.

[4] Chaurasiya, S. K., Biswas, A., Bandyopadhyay, P. K., Banerjee, A., & Banerjee, R. (2022). Metaheuristic load-balancing-based clustering technique in wireless sensor networks. *Wireless Communications and Mobile Computing*, *2022*, 1-21.

[5] Edla, D. R., Kongara, M. C., & Cheruku, R. (2019). A PSO based routing with novel fitness function for improving lifetime of WSNs. *Wireless Personal Communications*, *104*(1), 73-89.

[6] Wang, C., Liu, X., Hu, H., Han, Y., & Yao, M. (2020). Energy-efficient and load-balanced clustering routing protocol for wireless sensor networks using a chaotic genetic algorithm. *IEEE Access*, *8*, 158082-158096.

[7] Maheshwari, P., Sharma, A. K., & Verma, K. (2021). Energy efficient cluster based routing protocol for WSN using butterfly optimization algorithm and ant colony optimization. *Ad Hoc Networks*, *110*, 102317.

[8] Rajpoot, P., & Dwivedi, P. (2020). Optimized and load balanced clustering for wireless sensor networks to increase the lifetime of WSN using MADM approaches. *Wireless Networks*, *26*(1), 215-251.

[9] Rami Reddy, M., Ravi Chandra, M. L., Venkatramana, P., & Dilli, R. (2023). Energy-efficient cluster head selection in wireless sensor networks using an Enhanced grey wolf optimization algorithm. *Computers*, *12*(2), 35.

[10] Amutha, J., Sharma, S., & Sharma, S. K. (2022). An energy efficient cluster based hybrid optimization algorithm with static sink and mobile sink node for Wireless Sensor Networks. *Expert Systems with Applications*, *203*, 117334.

[11] Lipare, A., Reddy Edla, D., Cheruku, R., & Tripathi, D. (2020). GWO-GA based load balanced and energy efficient clustering approach for WSN. In *Smart Trends in Computing and Communications: Proceedings of SmartCom 2019* (pp. 287-295). Springer Singapore.

[12] Ludeña-Choez, J., Choquehuanca-Zevallos, J. J., & Mayhua-López, E. (2019). Sensor nodes fault detection for agricultural wireless sensor networks based on NMF. Computers and Electronics in Agriculture, 161, 214-224.

[13] Alshammari, H., Ghorbel, O., Aseeri, M., & Abid, M. (2018, June). Non-negative matrix factorization (NMF) for outlier detection in Wireless Sensor Networks. In *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)* (pp. 506-511). IEEE.

[14] Zhao, Z., Zhang, B., Zhang, B., Zhang, C., Wang, M., Yan, W., & Wang, F. (2021). A novel optimization method for WSN based on mixed matrix decomposition of NMF and 2-SVD-QR. Ad Hoc Networks, 115, 102454.

[15] Ghorbel, O., Alshammari, H., Aseeri, M., Khdhir, R., & Abid, M. (2020). Data Reduction Using NMF for Outlier Detection Method in Wireless Sensor Networks. In *Fourth International Congress on Information and Communication Technology: ICICT 2019, London, Volume 2* (pp. 23-30). Springer Singapore.

[16] Angel, M. A., & Jaya, T. (2022). An Enhanced Emperor Penguin Optimization Algorithm for Secure Energy Efficient Load Balancing in Wireless Sensor Networks. Wireless Personal Communications, 1-27.
[17] Rawat, P., & Chauhan, S. (2022). A novel cluster head selection and data aggregation protocol for heterogeneous wireless sensor network. Arabian Journal for Science and Engineering, 47(2), 1971-1986.
[18] Yousefpoor, E., Barati, H., & Barati, A. (2021). A hierarchical secure data aggregation method using the dragonfly algorithm in wireless sensor networks. Peer-to-Peer Networking and Applications, 14(4), 1917-1942.
[19] Tagne Fute, E., Kamdjou, H. M., El Amraoui, A., & Nzeukou, A. (2022). DDCA-WSN: A Distributed Data Compression and Aggregation Approach for Low Resources Wireless Sensors Networks. International Journal of Wireless Information Networks, 29(1), 80-92.
[20] Panimalar Kathiroli, Kanmani Selvadurai, "Energy efficient cluster head selection using Enhanced Sparrow Search Algorithm in Wireless Sensor Networks", Journal of King Saud University – Computer and Information Sciences 34 (2022) 8564–8575
[21] Muthukkumar, M. S., & Diwakaran, S. (2023). Efficient Load Balancing in WSN Using Quasi–oppositional Based Jaya Optimization with Cluster Head Selection. *International Journal of Computer Network and Information Security*, *13*(2), 85.
[22] KUMAR, M., & DIWAKARAN, S. (2023). Optimal Metaheuristic Sparrow Tuba Search Clustering for Load Balancing in Wireless Sensor Networks. *Adhoc & Sensor Wireless Networks*, *57*.
[23] Haseeb, K., Ud Din, I., Almogren, A., & Islam, N. (2020). An energy efficient and secure IoT-based WSN framework: An application to smart agriculture. *Sensors*, *20*(7), 2081.
[24] Edla, D. R., Lipare, A., Cheruku, R., & Kuppili, V. (2017). An efficient load balancing of gateways using Enhanced shuffled frog leaping algorithm and novel fitness function for WSNs. *IEEE Sensors Journal*, *17*(20), 6724-6733.
[25] Cervone, G., Franzese, P., & Keesee, A. P. (2010). Algorithm quasi-optimal (AQ) learning. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(2), 218-236.
[26] Osuna-Enciso, V., Cuevas, E., & Castañeda, B. M. (2022). A diversity metric for population-based metaheuristic algorithms. *Information Sciences*, *586*, 192-208.
[27] Trojovský, P., & Dehghani, M. (2023). A new bio-inspired metaheuristic algorithm for solving optimization problems based on walruses behavior. *Scientific Reports*, *13*(1), 8775.
[28] Aleksandra Urba´nczyk, Bartosz Nowak, Patryk Orzechowski, Jason H. Moore, Marek Kisiel-Dorohinicki, Aleksander Byrski "Socio-cognitive Evolution Strategies", DOI: 10.1007/978-3-030-77964-1_26
[29] Waintraub, M., PEREIRA, C. M. D. N. A., & SCHIRRU, R. (2007). The cellular particle swarm optimization algorithm.
[30] Gao, Y., Qian, C., Tao, Z., Zhou, H., Wu, J., & Yang, Y. (2020, December). Enhanced whale optimization algorithm via cellular automata. In *2020 IEEE International Conference on Progress in Informatics and Computing (PIC)* (pp. 34-39). IEEE.
[31] Zhu, C, Wu, S., Han, G., Shu, L., Wu, H, "A tree-cluster-based data-gathering algorithm for industrial WSNs with a mobile sink." IEEE Access,3, 381–396.
[32] R. Priyadarshi, L. Singh, Randheer and A. Singh, "A Novel HEED Protocol for Wireless Sensor Networks," 2018 5th International Conference on Signal Processing and Integrated Networks (SPIN), 2018, pp. 296-300, doi: 10.1109/SPIN.2018.8474286.
[33] Preeth, S.S., Dhanalakshmi, R., Kumar, R., Shakeel, P.M, "An adaptive fuzzy rule based energy efficient clustering and immune-inspired routing protocol for WSN-assisted IoT system". Journal of Ambient Intelligence and Humanized Computing, 1–13.

Authors' Profiles

Dr. Krishan Kumar, received his Bachelor of Engineering (B.E) degree in Electronics and Communication Engineering from Maharishi Dayanand University, Rohtak, Haryana, in 2003 and his Master of Technology (MTech) in VLSI Design and Embedded System from Guru Jambheshwar University of Science and Technology, Hisar Haryana, in 2008. At present, he is working as Assistant professor in Department of Electronics and Communication Engineering at Bhagat Phool Singh Mahila Vishwavidyalaya khanpur Kalan Sonipat, Haryana. He has completed his Ph.D. degree from School of Electronics and Communication Engineering from Shri Mata Vaishno Devi University Katra, Jammu &Kashmir India. His research interests are in the area of Artificial Intelligence, image and video processing, VLSI design and its implementation in VLSI. He has published more than twenty-five research papers in National/International/SCI and SCI-E journals and conferences.

Dr. Priyanka Anand is presently working as an Associate Professor in the Department of Electronics and Communication Engineering, Bhagat Phool Singh Mahila Vishwavidyalaya, Khanpur-Kalan, Sonipat, India. She has teaching experience of about 19 years in academic frontiers. She received her B.E. in Electrical Engineering in 2004, her MTech. in Electrical Engineering (Control & Instrumentation) in 2007, and subsequently her Ph.D. in Electrical Engineering in 2019. She has published more than 35 research papers in international (SCI/SCIE), national journals, conference proceedings, and book chapters. Her area of interest includes Renewable Energy, Hybrid Energy Systems, Soft Computing Techniques, Electric Vehicles, and Artificial Intelligence.

Dr. Rajni Mehra working as an Assistant Professor in the Department of Electronics and Communication Engineering at Deenbandhu Chhotu Ram University of Science and Technology. She has about 11 years of teaching experience. She received her B.Tech degree in Electronics and Communication Engineering with distinction, M.Tech. degree in Electronics and Communication Engineering with distinction from MDU university and PhD degree in Electronics and Communication from DCRUST University with distinction, Haryana. She has published 15 research papers in refereed international journals and 5 research papers in various international conferences. Her areas of research include Image Processing, Artificial Intelligence and Signal Processing etc.

How to cite this paper: Krishan Kumar, Priyanka Anand, Rajini Mehra, "Novel Data Compression and Aggregation Approach in WSN Using Enhanced Walrus Optimisation", International Journal of Computer Network and Information Security(IJCNIS), Vol.17, No.5, pp.15-30, 2025. DOI:10.5815/ijcnis.2025.05.02