RESEARCH PAPER

Analysis Of Spatial Correlation Based Multi-Hop Routing Scheme With Varying Nodal Energy Routing Agents For Wsns

Susheelkumar Sreedharan Panchikattil[1], Meenakshi R. Patil[2], Shilpa A N[3], Jayant Nandwalkar[4], Preeti Jain[5], Dr. Deepa R.[6], *Jayendra S. Jadhav[7]
[1] Dept. of ECE, CMR Institute of Technology, Bengaluru, INDIA
[2] Dept. of ECE, CMR Institute of Technology, Bengaluru, INDIA
[3] Selection Grade Lecturer, Dept. of Electronics and Communication Engineering, Government Polytechnic, Chamarajanagar, Dept. of Collegiate and Technical Education, Karnataka
[4] Dept. of Computer Engg., Datta Meghe College of Engineering, Airoli, Navi Mumbai, India
[5] Department of Artificial Intelligence & Data Science, Datta Meghe College of Engineering, Navi Mumbai
[6] Assoc. Prof., Dept. of AIML, AMC Engineering College, Bengaluru, INDIA
[7] Vishwakarma University, Pune, India
Corresponding Author: ***Jayendra S. Jadhav, Vishwakarma University, Pune, India**

Received: 20th Feb, 2026 | **Revised:** 4th Mar, 2026 | **Accepted:** 25th Mar, 2026 | **Available Online:** 10th Apr, 2026

ABSTRACT

Most applications of WSNs (Wireless Sensor Networks) demand an extended life expectancy, implementing improved and effective energy efficient measures. Clustering techniques are normally implemented to ensure improved energy utility and life-expectancy of sensor network. Multiple algorithms have been proposed with optimization techniques implemented for better energy efficiency and enhanced life-span of sensor networks. These algorithms arrange the sensing nodes into clusters and elect one of them as its leader. The data routing from the sensing nodes towards the sink may adopt either one-hop routing technique or multihop data routing technique. The functions of cluster head in a cluster are high energy consuming activities and hence its choice and the even distribution of cluster-head's (CH's) work-load among cluster members determines performance level of clustering algorithm. Similarly in multi-hop algorithms the probability of over burdening of routing nodes located near the sink is very high which hampers the performance level of routing algorithm. Hence clustering, CH selection and data flow path selection presents a computationally hard problem, which motivates us to explore metaheuristic algorithms to address the above issues for improved energy efficiency and network lifespan. Existing approaches use improved versions of standard Low Energy Adaptive Clustering Hierarchy (LEACH) protocol for clustering purposes and separate approach for data routing. Here we have proposed a spatial-correlation based clustering and data routing approach, wherein the computational complexities are reduced and only those cluster-head nodes are allowed to participate in the data-routing process which have an energy reserve higher than a pre-determined threshold. Here we have analysed the performance level of proposed algorithm against LEACH and its advanced variant in context of the throughput, network-Lifetime and the energy utility. Simulation results showcase enhancement of these quality-of-service metrics using our proposed algorithm as against the two standard algorithms. We have deeply investigated our multi-hop algorithm with varying parameters like node energy of the routing agent, spatial correlation constant considered for cluster formation. In addition, the dynamic nature of primary cluster-head and secondary cluster-head election also employs an effective data redundancy technique.

Keywords: Correlation, Clustering, Energy, Lifetime, Multi-hop, QoS metrics, Single-hop, Sink, Sensor Nodes, Throughput, Wireless Sensor Network.

How To Cite This Article: Panchikattil SS, Patil MR, Shilpa AN, Nandwalkar J, Jain P, Deepa R, Jadhav JS. Analysis of Spatial Correlation Based Multi-Hop Routing Scheme with Varying Nodal Energy Routing Agents for WSNs. Int J Drug Deliv Technol. 2026;16(28s):681-695. DOI: 10.25258/ijddt.16.28s.86

Source of support: Nil.

Conflict of interest: The authors declare no conflict of interest.

1 INTRODUCTION

WSN is recognized as one amongst the enabling technologies for today's IoT applications [1] cutting across defence and security sectors, agriculture sector and environment monitoring segments like seismological study. Many applications face constrained environments where conservation of energy is of utmost importance. WSN may be stationary or mobile in nature and it varies from application to application. Our focus of interest is aligned towards conserving energy for stationary WSN where high density of sensor nodes are scattered over a very large network area or terrain. These sensor nodes are empowered to communicate with other sensing nodes in its vicinity, and the terrain area under consideration restricts the energy replenishment of these nodes once their batteries are drained of power. Hence most algorithms [2-17] for data assimilation from such WSNs make use of clustering technique to conserve energy either implementing single hop or multiple hop structure for information data transfer from sensing end node to sink or base-station. Our model for simulation involves a large network area with a denser population of distributed sensing nodes. We shall analyse the simulation-model set-up using standard protocol LEACH [2, 3] and its advanced variant implementation called Enhanced LEACH [10] alongside our proposed scheme. The proposed algorithm uses a dynamic multi-hop and single hop algorithm, and care is accorded to facilitate uninterrupted data movement from source end to destination end even under network partitioning issues caused due to certain node failure or certain malicious behaviour exhibited by nodes. In our algorithm we have instilled selfishness behaviour among our cluster heads to act as routing agent by determining its residual energy after each round of data transfer. The simulation experiment with varying levels of minimum residual energy to act as routing agent facilitates the identification of best possible nodal residual energy value required for the agent to sustain a longer network lifetime. We have fine-tuned a very simplistic cluster grouping algorithm which makes use of spatial-correlation [18-29] between nodes to determine their cluster membership. The identification and assigned duties of cluster-head and sub cluster-head provides an add-on feature to conserve nodal energy, thereby prolonging its life-time. The redundancy in data transmission is addressed using the defined spatial-correlation parameter between cluster head and members of its cluster group and also using the defined spatial-correlation parameter between the cluster-heads spread across sensor network. The factors influencing the dynamic routing path selection in our multi-hop algorithm are the fixed minimalistic routing node energy and the next-hop distance befitting a chosen threshold distance. Next section-2 details the method of study, while section-3 investigates results with relevant discussion. Section-4 entails conclusion, while section-5 details the various abbreviations used and section-6 houses the declaration aspects with reference to this research and paper. The last section-7 assimilates the references used.

2 METHODS

Aim: To simulate an implementation using dynamic multi-hop data transfer structure towards the far-off sink utilizing the spatial-correlation between sensors to aid cluster grouping or formation and dynamic route selection for the data transfer with better results as against LEACH and Enhanced LEACH algorithms.

Design: Our simulation is based on the correlation model which revolves around the existing natural correlation between any two sensing nodes that is separated by a distance d. Now this distance d is normally doesnot exceed twice the sensing radius or range of two homogeneous sensing node and in case of heterogenous nodes, the separating distance d between the two heterogenous nodes is less than the summation of the sensing range or radius of the two individual nodes under consideration. Hence, for any two sensing node whose distance of separation is satisfying the above specified distance of separation criterion, we define a correlation that is expressed as a ratio of the overlapping coverage area between the two sensors and complete coverage area covered by both sensing nodes.

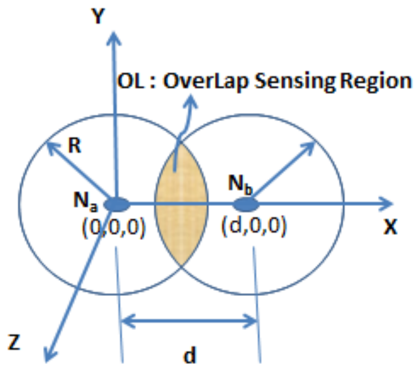

Fig. 1 Over-lapping region of sensing for two sensing nodes [20, 21, 26]

The Fig.1 depicts the above representation of two homogeneous sensing nodes, having uniform sensing-range or radius (R) and separated by distance d <2R and having overlapping area of coverage between the two sensing nodes defined as:

$$V_{OL} = \frac{\Pi}{12}(2\times R-d)^2(d+4\times R) \qquad (1)$$

And the total coverage area exhibited by the two homogeneous nodes is:

$$V_{Total} = \frac{8\,\Pi\times(R)^3}{3} - \frac{\Pi}{12}(2\times R-d)^2(d+4\times R) \qquad (2)$$

The correlation of the two nodes is expressed as a ratio of V_{OL} to V_{total} and is defined as:

$$\sigma = \frac{V_{OL}}{V_{Total}}$$

$$\sigma = \frac{(2\times R-d)^2(d+4\times R)}{32\times(R)^3-(2\times R-d)^2(d+4\times R)} \qquad (3)$$

Generalizing, we have:

$$\sigma = \begin{cases} \dfrac{(2\times R-d)^2\times(d+4\times R)}{32\times R^3-(2\times R-d)^2\times(d+4\times R)} & if \quad 0\le d \\ 0 & if \quad d\ge 2\times R \end{cases} \qquad (4)$$

Expression (4) is the mathematical representation for correlation model [20, 21, 26] used.

The Energy-flow model is taken from paper [14] and the standard parameter values for long range propagation and short-range propagation is taken and implemented in the simulation of the reference algorithms and our proposed algorithms. These parameter values are specified in table-1.

Setting of the study:

We have implemented a simulation for a Wireless Sensor Network Area of 200 * 200 metre^2. We have simulated a random distribution of 1000 sensor nodes in WSN area under consideration. Initial random distribution of the GPS-enabled sensing-nodes is kept the same for each iterations of study across multiple algorithms so that we can have a comparative analysis of results for identical sensor distribution. Also all the sensor nodes are energised with 0.5J from simulation point of view and real time nodes in real time implementation shall be energised accordingly depending on the application requirement and expected network life-span of the application.

The simulation of our improved methodology is effected using two distictive phases namely the centralized clustering process facilitated by the resource rich centralized sink and the distributed cluster head selection phase for the multi-hop data movement. The parameters chosen for simulation are the standard values as chosen for any wireless communication scenario specified in Table 1 and are detailed as follows. Each sensing node is initially uniformly charged to 0.5J and are assigned a uniform sensing radius range of 500 cms. Each message packet is assumed to be of 4000 bits with an energy expense of 0.5nJ/bit towards data aggregation at the sensor-node end.

The centralised sink is the final point of data aggregation in our data transmission process undertaken by each sensing node that are all initially uniformly charged or energized and stationary in position. Being location aware of all the distributed sensing nodes, it facilitates the implementtion of centralized clustering algorithms controled by the sink. This process of cluster grouping is initiated once in the sensor network's lifespan by sink which realys the cluster information along with its own location identity to all nodes, which ensures that all cluster member nodes are location-aware of all the co-member nodes. Correlation as defined by eq-4 is used for cluster formation by the centralized algorithm during each iteration of simulation. Here we have chosen the value of correlation criterion as 0.1 for iteration-one, 0.2 for iteration-two, 0.3 for iteration-three, 0.5 for iteration-four, 0.7 for iteration-five. Here we observe that as the value of correlation coefficient is increased from a value of 0.1 to 0.7, the size of the clusters decreases thereby reducing the cluster-members in each group while increasing the count of clusters or groups within the sensor network. Also with correlation coefficient value changing from 0.1 to 0.7, redundancy of data sensed by neighbouring sensor nodes also increases. Hence the choice of correlation coefficient should balance the requirement of data accuracy, network's lifespan and energy utility of network.

In each iterations, we have analyzed the simulations with five different levels of minimum energy fixated for cluster-heads to participate in multi-hop routing process. Hence for each iteration of correlation coeffient, we have obtained five separate sets of readings relevant to the five levels of minimum-energy fixated for CH participation in multi-hop routing process. The five levels of minimum-energy fixated for CH participation in multihop process are 0.4% of initial node energy i.e. 0.002J, 20% of initial node energy i.e. 0.1J, 40% of initial node energy i.e. 0.2J, 80% of initial node energy i.e. 0.4J and 100% of initial node energy i.e. 0.5J.

Thus the centralized sink groups the network of homogeneous nodes into separate clusters on the basis of initial defined correlation (σ) criterion fixed for the sensor network. These cluster groupings are maintained for the entire Network Life-time. Since the cluster details are relayed by the sink, the distributed nodes are location aware of the neighbouring nodes and keep a track of all the neighbouring nodes that are within a distance between 2SR and 5SR. This information helps in identifying the potential next hop node when it is in the process of data transfer. Hence the centralized clustering process that is decided upon by correlation criterion is a novel approach presented in this research paper. The main cluster representative (head) and the sub-cluster representative (head) choices as defined by the correlation criterion existing between cluster – head nodes and the self estimated figure of merit value by the node-end algorithm helps in enhancing the energy efficiency of the system. Here only the main-cluster representative (head) shall communicate with the next-hop routing node or the sink directly whereas any sub-cluster representative shall ony communicate with its associated main cluster representative (head), thereby reducing the energy expended towards next-hop data communication, which further adds to the novelty of our proposed approach.

Data packets that originate at any source node reach the sink follows the route as detailed further. Every cluster grouping has one of its member node function either as a cluster head or sub-cluster head defined by the estimated figure of merit value arrived at by node-end-distributed algorithm. All cluster representatives i.e the heads take-part in multihop data communication process as a next-hop routing node except for sub-cluster head which forwards its data to its correlated cluster representative or head as fixated by our algorithm. The cluster-member or the source node communicates its data-packet to associated cluster-head which also collates the received packets of data from each cluster-member and forwards collated data packets to next-hop routing node if it is designated as cluster head or otherwise in case of being designated as sub-CH, it will transmit collated data to correlated CH which then transfers its collated data to the next-hop routing node. In the absence of any potential next-hop routing node, head forwards data-packets directly to centralized-sink.

As discussed in earlier part, the implementation of our scheme is enacted in two phases i.e. the cluster grouping phase and data-transfer phase. First phase is controlled by centralized algorithm implemented by sink while the data communication i.e transfer phase is implemented at node-end using the distributed algorithm which facilitates multiple rounds of data routing to sink. At the begining of each round of packet communication, the cluster member node-end algorithm facilitates the election of head or representative of each cluster, designation of correlated-cluster representative (head) and sub-cluster representative (head) among elected cluster representatives (heads) during each data-transfer round. Each of the cluster representative (head) except the sub-cluster representatives (heads) also functions as next-hop routing node in the system for each round of data transfer. Criterion for election of the cluster representative or head and its further designation as correlated-cluster representative (head) and sub-cluster representative (head) adopts the following strategy. A figure-of-merit for each node in sensing network is estimated by distributed algorithm as defined by eq-5.

This Figure-of-merit for each node is arrived at based on its reserve or balance energy, distance of separation between the node and sink, and the cumulative data-flow path length of each member node to the potential cluster head in same cluster grouping governed by eq-5. This figure of Merit is advertised by each member node in cluster and node with the best figure of merit is seleced as cluster representative or head for that round of data transfer which is further advertised by CH to all neighbouring cluster heads within a reach of 5 times SRs. This is considered to be the period-I of the data transfer phase during each rounds of data transfer. The cluster head that satisfies a spatial correlation factor that is greater than or equal to 0.1 with its neighbouring cluster head are either redesignated as Correlated CH or sub CH based on the comparative values of their individual figure-of-merit. The cluster representative or head with the highest value of figure-of-merit acts as the main cluster representative or head and the other cluster representatives or heads that fulfills minimum correlation-value equal-to 0.1 are designated as the sub-cluster representative or head which shall forward its assimilated cluster message-packets to main correlated cluster representative or head. Sub-cluster representatives or heads can never function as the routing node while they can only forward their aggregated data to their correlated-cluster representative or head which also doubles up as a routing agent too. The next-hop routing node can be any of the cluster representatives or heads, except the sub-cluster representatives or heads, that is at a distance of 2SR to 5SR from the relaying cluster head or representative. The check on the redundancy of data and effective energy utilization is implemented within the algorithm itself by making all the nodes that fulfill a cluster formation minimum correlation factor of $\sigma = 0.1$ with its CH (cluster representative or head) or the correlated-cluster representative or head, dormant for that data transfer round. All the cluster representatives or heads except the sub-cluster head or representative also finds the best distance based next-hop routing node so as to channelize communication from node end to sink. In the absence of any such further routing node in the viscinity, the relaying CH node relays data directly to sink. Sink keeps a track of all the live nodes and dead nodes in the system, by rendereing any nodein the network as inactive if that node does not participate in five successive data transfer rounds. The centralized algorithm also esimates the time required for the data to be received from the farthest cluster during each data transfer rounds. This estimation helps the sink to dictate the start of next round of data transfer across the network. This algorithm is implemented using MATLAB 2016, wherein the Spatial Correlation factor is varied between 0 to 1 and its performance is benchmarked against Standard LEACH and its enhanced variant namely Enhanced LEACH in terms of the three quality of service metrics i.e. Network Lifetime, Energy Utilization factor and throughput achieved by the system.

ALGORITHMS

Simulation is carried out in two distinctive phases namely the centralized clustering process facilitated by the resource rich centralized sink and the distributed cluster head selection phase for the multi-hop data movement. Cluster formation centralized algorithm implemented at the sink end is adopted from paper [20-21]. The cluster groupings are finalized at the sink end and the cluster defining details are published by sink node to distributed sensor-nodes, so that each node is location aware of the other cluster members in the group as well as the sink location detail. On the contrary, the distributed algorithm that is run at each and every sensor node facilitates the cluster representative (head) anointment for each cluster, designating appropriate cluster representative (head) as the correlated-cluster representative (head), or sub-cluster representative (head), next hop route selection and lastly effecting the data transfer to the sink.

Fig. 2 showcases the flowchart of Algorithm-1 which is our centralized cluster formation algorithm at sink end with following assumptions:

1. Sink is resource enriched from the perspective of processing power, storage and energy requiement.
2. It is abreast with node specifics like their identities, their locations, IDs and sensing range.
3. It assumes same node sensing radius (5 units) for every node in the sensor network

Algorithm-1:

- Sensing range of 5 unit radii is assumed
- Evaluate Sensing-region-Overlap between adjacent nodes (say node-1 &2)

- Estimate spatial correlation between them i.e. σ_{12}
- If $\sigma_{12} > \sigma_{chosen\ for\ Algo}$, then flag node-2 as grouped with node-1 else flag node-2 as not grouped
- Repeat above steps keeping node-1 as the central node and sequentially replacing node-2 with the remaining nodes of the network one by one to form the set of grouped-nodes with node-1 identified as cluster-1
- Similarly, select next ungrouped node as centre node and repeat from steps-2 to form multiple clusters

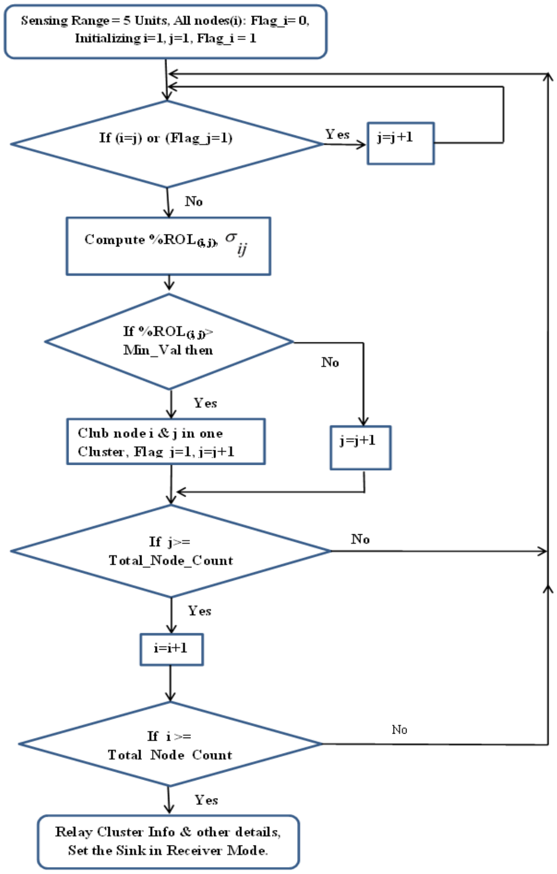

Fig. 2 Centralized Sink-end Algorithm 1[20,21]

Fig. 3 showcases second algorithm (flowchart) which is distributed and implemented at node end. This algorithm initializes cluster representative election and enables the data flow phase.

Algorithm-2:

- Estimate residual or balance energy of node E_{BAL_N}
- Estimate the euclidean path length of other cluster-member nodes from itself and the path length to sink
- Estimate the total intra-cluster summative distance to be traversed for message packets from other cluster-members to reach it i.e. $Intra\_CLST\_DIST$
- Estimate the Figure of Merit (FM) of node
- Broadcast FM value to the cluster and neighbouring cluster heads
- Highest valued FM to be group or Cluster representative
- Re-designate as main or associate-Cluster representative (head) or sub-cluster representative (head) or normal cluster representative (head)
- Identify the next-hop routing node
- Start the data transfer phase and participate in data routing so that the data packets multi-hops and reaches the destination sink.
- Sink terminates the data transfer phase after the prefixed duration estimated by sink
- Culmination of current Round and then repeat from step-1 for subsequent rounds

Distributed algorithm is responsible for the estimation of Figure-of-Merit (FM). The Figure of Merit (FM) value is arrived at by the expression [20-21]:

$$FM = \left(\frac{W_{D2S} \times E_{BAL\_N} + W_{Intra\_Clust}\left(\frac{1}{Intra\_CLST\_DIST^{2}}\right)}{E_{BAL\_N} + \left(\frac{1}{Intra\_CLST\_DIST^{2}}\right)} \right) \quad (5)$$

Eq-5 determines the best cluster member to take-on role of the group representative in any cluster. Also, this parameter facilitates the selection of sub cluster head between two neighboring cluster-heads having a correlation coefficient of minimum 0.1. Here W_{D2S}, is the weight that is proportional to the distance from sink to node and W_{Intra_Clust} is the weight that is relative to the total intra-cluster propagation path length to be covered incase node becomes Cluste-head. Further E_{BAL_N}: balance energy with the sensor node at the start of data transfer round and $Intra\_CLST\_DIST$ is the sum total of the length of intra cluster distance of data propagation that is effectively traversed when the members forward their data to its cluster head. Both weights are defined as:

$$W_{D2S} = \left(\frac{PD_{S2N}}{PD_{S2N} + \text{Intra\_CLST\_DIST}}\right) \qquad (6)$$

,

$$W_{Intra\_Clust} = \left(\frac{Intra\_CLST\_DIST}{PD_{S2N} + Intra\_CLST\_DIST}\right) \qquad (7)$$

where PD_{S2N} is the propagation path from sink to node under consideration and $Intra\_CLST\_DIST$ is sum total of the intra-cluster propagation path length that needs to be traversed to forward data from CM to CH. W_{D2S} and W_{Intra_Clust} dynamically estimates the value of FM as per status of the node's energy, intra-cluster propagation path length and the propagation path length from candidate cluster-head to sink, during the beginning of data transfer round and thereby facilitates the election of cluster-head from the members of the cluster group. Distributed network's node-end algorithm helps in the above determination along with, identifying the sub-cluster heads that shall be forwarding its aggregated data to its correlated cluster head. The forward routing path is also facilitated by this distributed algorithm with the help of a TDMA schedule decided by the forwarding CH to all back CHs that are in vicinity range of 2SR to 5SR. A separate TDMA Schedule is maintained by each Custer Head to aggregate data from each of its cluster members and the sub-cluster heads. Therefore during each data transfer round, initially the cluster head, correlated cluster head and sub cluster head are determined in period-I and the TDMA schedules are forwarded by the correlated cluster heads to its associated cluster members and sub-cluster head. Similarly, the TDMA schedules are forwarded by sub-cluster heads to its associated cluster members. The forward routing cluster heads also maintains a separate TDMA Schedule for all its associated back-cluster heads. Period – II encompasses all the TDMA scheduling as briefed earlier. Period-III reflects the actual data transfer wherein the TDMA scheduling for back-cluster head is repeatedly used for data transfer until the sink alerts the end of data transfer round. After which, the routing cluster heads drops the data, if any still pending to be relayed and initiates itself to a fresh data transfer round along with other member nodes. Sink also determines the length of each data transfer round based on total number of active nodes, clusters formed, overall estimation of propagation path length considering the longest path and buffer time included so as to allow the data transfer from the farthest cluster in the network. The sink analyzes the received data and identifies with the respective cluster and respective sensor node that has relayed the sensed data. The sink also maintains a record of all the active nodes in the system and declares any node as inactive, incase there is no data transfer from any network-end node for last five consecutive data transfer rounds. This information on the identity of dead nodes is also relayed by sink when it alerts fresh data transfer round, so that each active network-end node is also aware of its dead co-cluster members.

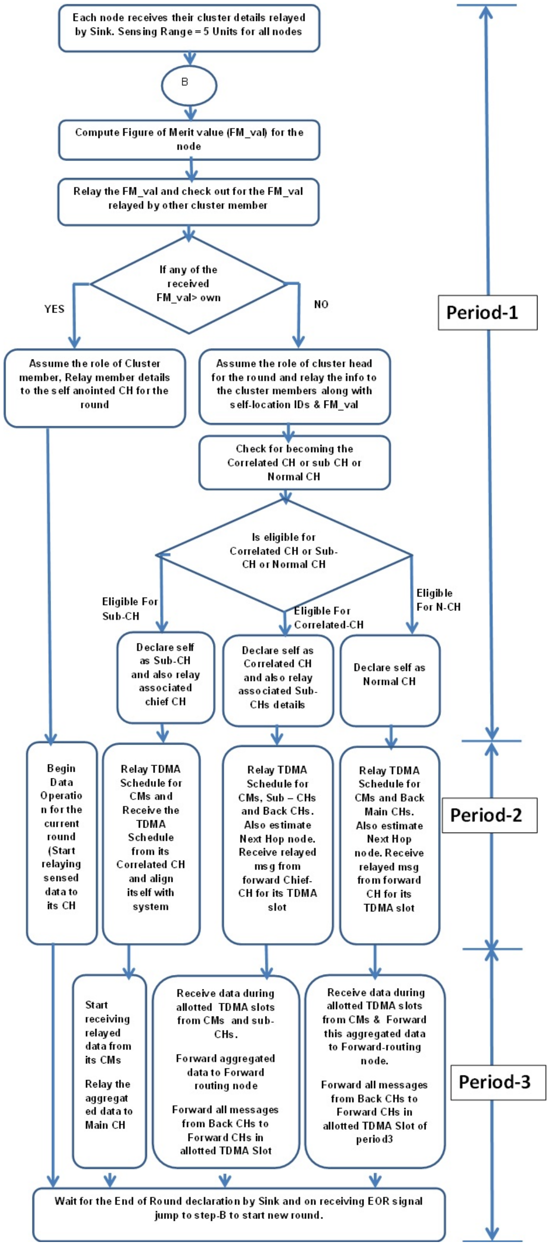

Fig. 3 Distributed Algorithm (Algorithm-2)

3 RESULTS AND DISCUSSION

Many of the related works are based on the working of standard LEACH protocol by incorporating multiple enhancements in the algorithms used. In paper [6], authors R.Sinde et.al. (2020) have introduced E2S-DRL algorithm to effect better lifetime and to reduce delay. Under this protocol, they have used Zone based clustering scheme involving the hybrid PSO (Particle-Swarm-Optimization) and AP (Affinity-Propagation) algorithms to facilitate data aggregation with reduced energy consumption. Further they have used ACO (Ant-Colony-Optimization) and FFA (Fire-Fly) Algorithm to effect a routing algorithm with reduced delay. They have effectively controlled the duty cycle of each node by altering or adaptively selecting the scheduled mode for each node with the help of Deep Reinforcement Learning (DRL) algorithm. Their simulation results have showcased improved energy consumption, throughput with reduced delay. The algorithms simulated are complex ones which necessitate a complex computational processing at each node.

In standard LEACH protocol, we know that the cluster-head is elected using a randomized probabilistic approach and each round of data transmission involves two distinct phases distinguished as set up or initializing and data movement or transfer phases. First phase oversees CH (head) election and formation of clusters while the second phase facilitates data movement from sensing-node to CH head and transfer of aggregated data from CH to central-sink. Drawbacks of randomized selection of CHs during each and every round of data transfer in LEACH is corrected by incorporating a weighted CH election decided on basis of leftover-energy and continuation of its role as cluster-head into subsequent rounds of data transfer till its residual energy falls below a threshold value in MODLEACH [11]. The authors of MODLEACH have incorporated a two-level type of amplification during data transmission depending on whether it's an intra-cluster communication or a communication between the cluster-head and sink.

In [12], the authors have showcased a dynamic data aggregation scheme by activating the nodes near the relevant event area under consideration. The authors have worked on estimating the optimal number of nodes to be activated for the given cluster area to minimize the distortion in the sensed parameter of the given event area of interest. Hence the authors have successfully increased the reliability of the information sensed from the event area. They have used the spatial correlation to dynamically select the nodes responsible for sensing and contribute to the information, in effect reducing the redundancy. In [19], Zhao et al. (2018) have tried to bring about a balancing attempt to utilize the energy in an optimum way between inter-cluster and intra-cluster activities. They have implemented node dormancy scheme that has added further energy saving. But it is observed that their scheme of implementation is best suited for small sized network area since the scheme involves only two or three hop data transfer to sink. In [10], the authors have worked around the standard LEACH protocol and have come up with a modified or enhanced version by considering both the intra-cluster distance of communication and distance from CH to sink to segment sensor network into clusters. The simulation of their algorithm has showcased an increased life-span for the sensor network functioning as against standard LEACH protocol. But even this implementation of the enhanced LEACH is limited by the smaller size of sensor network area since the data transfer across a larger network area cannot be efficient using a two-hop or three hop structures.

The hierarchical chimp optimization (HChOA) introduced in paper [30] uses a metaheuristic clustering and routing multi-hop algorithm for under water sensor networks. He *et al*. have introduced a common efficient and multi-hop clustering algorithm by integrating the two processes to yield better results against benchmarks like LEACH, MPSO, PSO, TEEN, IPSO-GWO for metrics like network lifetime and energy utilization. EECRU algorithm introduced by Mohammed Ali *et al*. addresses the clustering process in under water sensor networks using special nodes that act as header nodes and in the absence of header nodes make use of neighbour node method [31]. Their technique also implements a CDMA concept to get information from special and other nodes. The special node placement or location impacts the overall energy consumption which is not considered in the algorithm.

Yang Yang et al. have proposed an hybrid Chimp Optimization based ChOA-HGS algorithm [32]. In this paper, authors have use ChOA for selecting the cluster-heads and make the clusters while HGS shows the best path for data routing the data flow. Their proposed algorithm have yielded better results against benchmarks like PSO, IPSO_GWO, TEEN, MPSO and LEACH. Kisheh et al. have explored an efficient way using AI based ant-lion optimizer to forecast electroencephalogram data [33,34]. But the complex computations restricts its usage in energy constrained networks.

Houman et al. [35] have used grasshoper optimization algorithm (GOA) in an hybrid classifier for calssifying sonar related big data. The results of their simulation suggests GOA can estimate the correct boundary between exploration and exploitation phases which enables it to come out of local optima and move towards global optima for resolving multi dimensional issues. The complexities involved in the implementation of artificial neural networks and connected GOA algorithm would restrict their usage in energy constrained sensor networks.

Most of these algorithms are updations and enhancements of standard LEACH and hence we have chosen standard LEACH and Enhanced LEACH as the benchmark against which we compare our proposed algorithms performance. Here we have chosen three quality of service metrics for comparison purpose namely throughput, network lifetime and energy utility factor. Throughput has been assumed as the aggregate of message packets reaching centralized sink during the sensor network's lifetime. Graphically it is represented here as the incremental count of message packets reaching destination sink with increasing rounds of data transfer.

Network Lifetime has been assumed as the life span of the network till 70% network-nodes is dead. Graphically, we have represented the incremental count of dead-nodes versus the incremental rounds of data transfer. Energy utility factor gives the detail of energy being consumed in the network during network's lifespan. It has been represented as the balance energy leftover in system network with each incremental round of data movement. We analyse the energy efficiency of each system by checking the balance-energy in the sensor network after fifteen-hundred (1500) data transfer rounds since it is seen that all nodes in the sensor network is completely depleted of their entire energy by around fifteen hundred (1500) rounds in case of LEACH or Enhanced LEACH. The balance energy of the network refers to the cumulative sum of balance individual alive node's energy.

Table 1: Simulation Parameters

No	Parameters	Value
1	Distributed number of Sensing or Network-end Node	1000
2	Network Span (Area)	200*200 sq.m
3	Network-end node's starting Energy	0.5 J
4	Individual node's sensing range	500 cm

5	Energy required for data aggregation	5 nJ/bit
6	Message packet length	4000 bits
7	Communication Energy required (reception/ transmission)	50 nJ/bit
8	Free-Space factor: d<d0 (cross-over-distance)	10 nJ/bit/m2
9	Multi-path propagation factor for d>d0	0.0013 pJ/bit/m4

Implementation is based on the simulation using standard parameters adopted from paper [28-29] which are depicted in the above table and were detailed in the methodology section above. The simulation showcases the result for a wide range of correlation coefficient factors spanning between values from 0.1 to 0.7, implemented for our proposed multi-hop algorithm. Each implementation with specific correlation coefficients checks out the results for five distinct varying values of minimum node energy required to support the functioning of node as routing agent. The multiple instances least-energy-requirement to participate in routing process for any cluster head other than sub cluster head is kept at 0.4% of initial node-energy, 20%, 40%, 80% and 100% of the initial node energy. These results are verified against standard LEACH protocol and Enhanced LEACH protocol and are tabulated in table-2. The same is plotted in fig. 4 (a), 4(b), 4(c) to fig. 8(a), 8(b), 8(c) for visualization. Fig. 4(a), 4(b) and 4(c) showcases the performance of various instances of our proposed algorithm with correlation coefficient of $\sigma = 0.7$ versus the standard LEACH and Enhanced LEACH algorithm. Fig. 5(a), 5(b) and 5(c) showcases the performance of various instances of our proposed algorithm with correlation coefficient of $\sigma = 0.5$ versus the standard LEACH and Enhanced LEACH algorithm. Fig. 6(a), 6(b) and 6(c) showcases the performance of various instances of our proposed algorithm with correlation coefficient of $\sigma = 0.3$ versus the standard LEACH and Enhanced LEACH algorithm. Fig. 7(a), 7(b) and 7(c) showcases the performance of various instances of our improved algorithm with $\sigma = 0.2$ versus the standard LEACH and Enhanced LEACH algorithm. Fig. 8(a), 8(b) and 8(c) showcases the performance of our algorithm's various instances with correlation coefficient $\sigma = 0.1$ versus the standard LEACH and Enhanced LEACH algorithm.

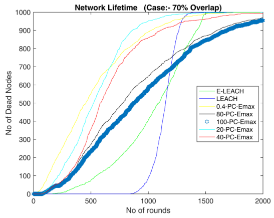

Fig. 4 (a) Network Lifetime ($\sigma = 0.7$)

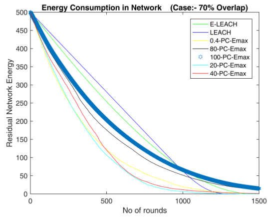

Fig. 4(b) Energy Utility ($\sigma = 0.7$)

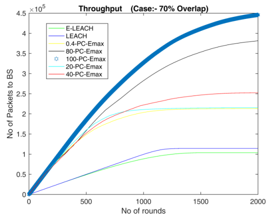

Fig. 4 (c) System Throughput ($\sigma = 0.7$)

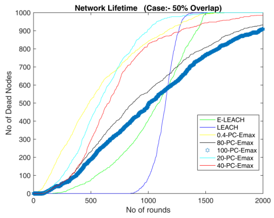

Fig. 5(a) Network Lifetime ($\sigma = 0.5$)

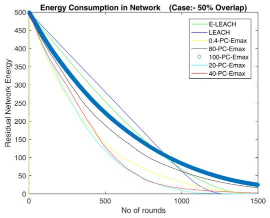

Fig. 5(b) Energy Utility ($\sigma = 0.5$)

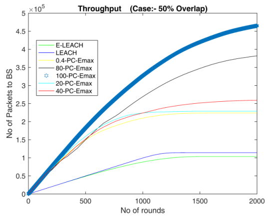

Fig. 5(c) System Throughput ($\sigma = 0.5$

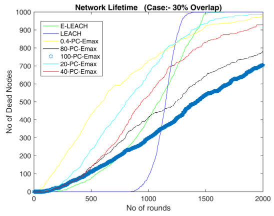

Fig. 6(a) Network Lifetime ($\sigma = 0.3$)

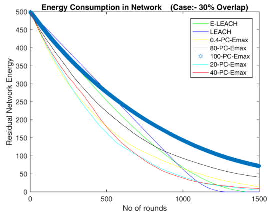

Fig. 6(b) Energy Utility ($\sigma = 0.3$)

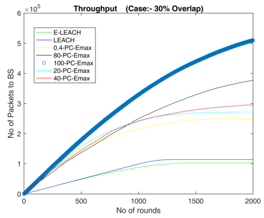

Fig. 6(c) System Throughput ($\sigma = 0.3$)

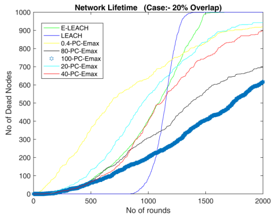

Fig. 7(a) Network Lifetime ($\sigma = 0.2$)

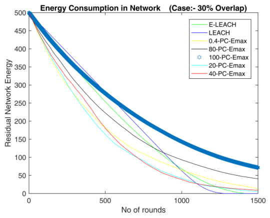

Fig. 7(b) Energy Utility ($\sigma = 0.2$)

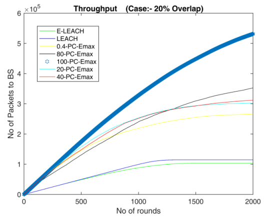

Fig. 7(c) System Throughput ($\sigma = 0.2$)

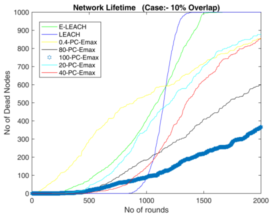

Fig. 8(a) Network Lifetime (σ = 0.1)

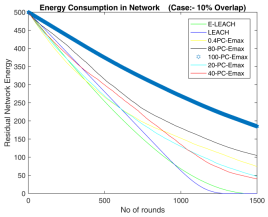

Fig. 8(b) Energy Utility (σ = 0.1)

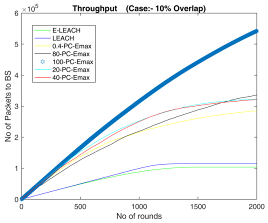

Fig. 8(c) System Throughput (σ = 0.1)

Table 2 Proposed Algorithm with Coefficient correlations and its corresponding results against benchmarked standard-LEACH and Enhanced LEACH Algorithms

Protocols : Recent and Proposed Algorithms	No of Dead Nodes								Throughput at 70% Node Deaths or at Round no. 2000	Residual Network Energy after 1500 rounds of Data Transfer
	100	200	300	400	500	600	700	800		
	Round of Data Transfer wherein the above Number of Dead Nodes are observed									
Column ID	1	2	3	4	5	6	7	8	9	10
LEACH	1011	1065	1097	1124	1152	1172	1198	1227	112795	0
EN_LEACH	546	750	893	996	1091	1182	1252	1322	100009	0
Correlation										
Coefficient (σ) /Overlap % with min routing Energy for CHs										
0.1 / 10PC_0.4%_Emax	237	404	508	661	885	1217	1454	1780	257825	74.53
0.1 /10PC_20%_Emax	659	810	916	1042	1205	1372	1493	1680	305422	47.06
0.1 /10PC_40%_Emax	917	1077	1179	1272	1348	1459	1621	1838	308290	40.48
0.1 /10PC_80%_Emax	768	1040	1288	1489	1712	1998	>2000	>2000	336230	104.89
0.1 /10PC_100%_Emax	1064	1410	1811	>2000	>2000	>2000	>2000	>2000	542638	185.19
0.2 /20PC_0.4%_Emax	197	321	445	546	671	875	1130	1368	231426	33.44

0.2 /20PC_20%_Emax	466	627	745	848	948	1085	1241	1421	278972	23.74
0.2 /20PC_40%_Emax	590	798	892	1000	1113	1248	1395	1609	287724	18.13
0.2 /20PC_80%_Emax	568	780	971	1211	1370	1569	>2000	>2000	352825	58.13
0.2 /20PC_100%_Emax	682	996	1272	1475	1734	1964	>2000	>2000	531498	102.51
0.3 /30PC_0.4%_Emax	161	269	382	470	595	764	952	1184	211740	13.08
0.3 /30PC_20%_Emax	332	466	576	699	803	947	1050	1161	248286	5.37
0.3 /30PC_40%_Emax	445	596	741	831	941	1075	1205	1419	260441	9.24
0.3 /30PC_80%_Emax	431	618	798	981	1238	1409	1677	>2000	377568	40.85
0.3 /30PC_100%_Emax	471	747	1005	1236	1414	1698	1974	>2000	510162	71.78
0.5 /50PC_0.4%_Emax	139	217	312	388	484	587	750	919	187308	0
0.5 /50PC_20%_Emax	256	354	425	491	576	654	737	857	202943	0
0.5 /50PC_40%_Emax	331	439	503	573	656	738	816	959	207062	2.80
0.5 /50PC_80%_Emax	328	456	591	721	866	1060	1247	1474	314122	16.79
0.5 /50PC_100%_Emax	374	518	668	850	1007	1202	1363	1605	400748	24.94
0.7 /70PC_0.4%_Emax	126	205	288	370	457	546	674	816	175487	0
0.7 /70PC_20%_Emax	212	311	395	454	524	598	670	750	191025	0
0.7 /70PC_40%_Emax	293	399	467	532	582	662	740	876	194921	1.05
0.7 /70PC_80%_Emax	317	413	531	646	777	909	1094	1315	300201	10.63
0.7 /70PC_100%_Emax	324	448	580	735	876	1032	1217	1369	372316	14.93

The tabulated observations lead to the following inferences:

Observation and Inference-1: In the standard protocol LEACH, it is observed that the Network lifetime defined by 70% node deaths is seen to take effect at data transfer round number 1198 and during this life time, it has supported a throughput of 1,12,795 packet delivery to sink. Enhanced LEACH protocol puts up an almost similar scenario wherein 70% node-deaths is reflected at around data transfer round number 1252 but with an effective throughput of 1,00,009 packets of message being delivered to the centralized sink in its life time.

By around 1500 data transfer rounds, leftover or balance energy in sensor-network system for benchmarked algorithms namely LEACH and its enhanced-variant approaches zero.

Observation and Inference-2:

In contrast, our Proposed algorithm have been simulated for various correlation coefficients. To analyze our algorithms in a simplified manner, we first look at the Network energy point of view for all the iterations of correlation coefficients, with five different levels of minimum-Energy fixated required for participation in multi-hop data routing by the main cluster-heads in each iteration. Here from the last column of table-2, we observe that the balance energy for each group of iterations after 1500 rounds of data transfer goes on decreasing as the correlation coefficient is varied from 0.1 to 0.7. Hence the best results are observed in the iterative group with correlation coefficients of 0.1 and 0.2, with coefficient value of 0.1 better placed than 0.2.

Also, within each of the iterative group readings in table-2, it is seen that the leftover-energy with sensor-network system after 1500 data transfer rounds increases as the level of minimum-fixated-energy for multi-hop data routing participation is increased. Thus, from an energy conservation point of view, our algorithm conserves more energy and is at its best at correlation coefficient value of 0.1 followed by the 2nd best at correlation coefficient value of 0.2

Observation and Inference-3:

The column with ID-9 of table-2 presents the details of the throughput in terms of the aggregate count of message packets delivered to centralized sink during network's lifespan. Here its observed that the throughput in the case of our algorithm with minimum-Energy-fiixated for data routing participation equated to 0.4% of initial node energy i.e. 0.002J is almost twice the throughput encountered in case of the benchmarked algorithms. Also the throughput in case of our algorithm is seen to increase with the increase of minimum-Energy-fiixated for data routing participation from 0.4% to 100% of initial node energy.

4 CONCLUSIONS

From the above discussion over the tabulated observations and the results, we conclude that our algorithm with $\sigma = 0.1$ and $\sigma = 0.2$ shows improved conservation of energy which is evident from column-10 of table-2, thereby enhancing the lifetime of the network. Hence, considering the aspect of energy consumed in network, performance of our algorithm is best for lower values of correlation coefficients and almost gives an improved energy utility on varying the minimum-Energy-fixated for data routing participation from 0.4% to 100% of the maximum initial energy available with each node at the beginning especially with a cluster formation correlation coefficient (σ) of 0.1 and 0.2 in comparison to results obtained in the bench-marked algorithms.

Now from the perspective of throughput, it is observed from colum-9 of table-2 that the total message packets delivered in case of improved algorithm with the $\sigma = 0.1$ and $\sigma = 0.2$ at a minimum-energy-fixated level of 0.4% of initial node energy is almost double than the observed results for our bench-marked algorithms. Further analysis of data specified in column-9 of table-2 reveals that within each iterative group represented by $\sigma = 0.1$ and $\sigma = 0.2$, the throughput for the improved algorithm increases from twice to five times the observed results in case of bench-marked algorithms as we raise the minimum-Energy-fixated levels for data routing participation from 0.4% to 100% of initial node energy.

From the perspective of network lifetime also, the performance of our algorithm is best for lower values of correlation coefficients which is evident from the readings noted in column-7 of table-2. We observe an improved network life of additional 20% to 50% and above, on varying the minimum-Energy-fixated for data routing participation from 0.4% to 100% of the initial node energy available with each node at the beginning especially with a cluster formation correlation coefficient (σ) of 0.1 and 0.2 in comparison to results obtained in standard-LEACH and Enhanced-LEACH algorithms.

We may therefore conclude that our algorithm with correlation coefficient of 0.1 and 0.2 with five shades of minimum required energy values for routing participation performs better that standard LEACH and Enhanced LEACH.

5 LIST OF ABBREVIATIONS

LEACH: Low Energy Adaptive Clustering Hierarchy
IoT: Internet of Things
QoS: Quality of Service
E2S-DRL: Energy-Efficient Scheduling using the Deep Reinforcement Learning
PSO: Particle-Swarm-Optimization
AP: Affinity-Propagation
FFA: Fire-Fly Algorithm
HChOA: Hierarchical Chimp Optimization Algorithm
IPSO-GWO: Improved Particle Swarm Optimization - Grey Wolf Optimization
EECRU: Energy-Eficient and Cluster-based Routing protocol

CDMA: Code Division Multiple Access

ChOA-HGS: Chimp Optimization and Hunger Games Search

GOA: Grasshoper Optimization Algorithm

E_{BAL_N}: Residual or Balance Energy of node

$Intra\_CLST\_DIST$: Intra-Cluster summative Distance

FM: Figure of Merit

PD_{S2N}: Propagation path Distance from Sink to Node

W_{D2S}: Weight proportional to Distance from Sink to Node

W_{Intra_Clust}: Weight proportional to total Intra-Cluster propagation path length

SR: Sensing Range

TDMA: Time Division Multiple Access

6 DECLARATIONS

- Data and material Availability: All data used in the paper shall be made available upon request.
- Competing interests: The authors declare that they have no competing interests.
- Funding: Not Applicable.
- Authors' contributions: SSP analyzed the concept of spatial correlation existing between sensing nodes, arrived at the mathematical representation of the same and carried out the simulation for the proposed algorithm with varying correlation coefficients using MATLAB @2016. KT carried out the simulation of the existing algorithms namely LEACH and Enhanced LEACH and interpreted its results. MP was a major contributor in writing the manuscript and interpreting the observed results in graphical formats. DR carried out the comparative analysis of the observed results for the proposed algorithm and the existing algorithms. SS focused on the development of the improved centralized algorithm and its implementation in MATLAB simulation. SV revised and updated the distributed algorithm and carried out its implementation in simulation.
- Acknowledgements: Not Applicable

7 REFERENCES

[1] Sohraby K., Minoli D. and Znati T., WSNs: Technology, Protocols and Applications, John Wiley and Sons, Hoboken, New Jersey, pp. 38-71, 2007.

[2] Heinzelman W., Chandrakasan A. and Balakrishnan H., 2000, “Energy-Efficient Communication-Protocol for Wireless Microsensor Networks”, 33rd Hawaii Int Conf on Sys Sciences, 2000.

[3] Heinzelman W., Chandrakasan A. and Balakrishnan H., “An Application-Specific Protocol Architecture for Wireless Microsensor Networks”, IEEE Trans Wireless Commun, vol. 1(4), 2002, doi: 10.1109/TWC.2002.8.

[4] Farouk F, Rizk and W. Zaki, "Multi-level Stable and Energy-efficient Clustering Protocol in Heterogeneous Wireless Sensor Networks", *Wirel. Sens. Syst. IET*, vol. 4, no. 4, pp. 159-169, 2014.

[5] Shah, Abbas Z, Abbas, Muhammad F, Hussien, Baker, “An Innovative Clustering Hierarchical-Protocol for Data Collection from Remote WSNs Based IOT Applications”, Sensors (Basel). 2023 Jun 19;23(12):5728. PMID: 37420893; PMCID: PMC10305020.

[6] Sinde, Begum, Njau, Kaijage, “Refining Network-Lifetime of WSN using Energy-Efficient Clustering and DRL-based Sleep-Scheduling”, MDPI Sensors, 2020.

[7] Ouafa, K. Salah, L. Jalal, El. Said, “Recent Advances of Hierarchical Routing Protocols for Ad-Hoc and Wireless Sensor Networks: A Literature Survey,” BİLİŞİM TEKNOLOJİLERİ DERGİSİ, CİLT: 9, SAYI: 2, MAYIS 2016.

[8] Golam R., Kabir M., M. Sajadur, U. Sheik, “Cluster Based Hierarchical Routing Protocol For Wireless Sensor Network”, IJCNS, vol. 2/5, 2010.

[9] M. Alom, Hosan A, K. Palab, “Improved Zonal Stable Election Protocol (IZ-SEP) for hierarchical clustering in heterogeneous wireless sensor networks”, e-Prime-Advances in Electrical Engineering, Electronics and Energy, Vol. 2, 2022.

[10] Amer, Salem, Noor, “Enhanced-LEACH Protocol for Increasing a Lifetime of WSNs”, Personal and Ubiquitous Computing, pp .901–07, 2019, @Springer-Verlag London Ltd.

[11] D. Mahmood, N. Javaid, S. Mahmood, S. Qureshi, A.M. Memon, T. Zaman, “MODLEACH: a variant of LEACH for WSNs”, eighth international conferenceon broadband, wireless computing, Commun. Appl. (2013) 158–163, http://dx.doi.org/10.1109/BWCCA.2013.34.

[12] Floriano DE RANGO, Nunzia PALMIERI, Simona RANIERI, “Spatial Correlation Based LowEnergyAware Clustering (LEACH) ina Wireless Sensor Networks”, Information and Communication-Technologies and Services Volume: 13 Number: 4, 2015.

[13] Arora V., Sharma V. and Sachdeva M., “A Survey on LEACH and other Routing Protocols in Wireless Sensor Network”, Elsevier Optik Journal, Optik 127, pp. 6590–6600, 2016.

[14] F. Comeau and N. Aslam, "Analysis of LEACH Energy Parameters", Workshop on Em-Se-Ns, ELSEVIER, Science Direct, pp. 933-938, 2011.

[15] Tamilarasi N. and Parvathi K., "Extension of Network Lifetime using Fuzzy C-Means, Model and Cluster Hierarchy Concept", International Journal of Recent Trends in Engineering & Research, vol. 2(8), pp. 302 - 308, 2016.

[16] Toor, Jain, "Energy aware cluster based multi-hop energy-efficient routing protocol using multiple mobile-nodes in WSNs," AEU – Int. Journal of Electronics and Communications, vol. 102, pp. 41–53, 2019.

[17] R. Yarinezhad, "Reducing delay and prolonging the lifetime of WSN using efficient routing-protocol based on mobile-sink and virtual-infrastructure", Ad-Hoc Networks, vol. 84, pp. 42–55, 2019.

[18] Yu.Z, J. Qian, Liu S., W. Chong, "A Novel Spatial Correlation-Based Clustering Protocol in Wireless Sensor Network", 4th International Conference on Wireless Communications, Networking and Mobile Computing, Dalian, China, 2008, pp. 1-4, doi: 10.1109/WiCom.2008.965.

[19] Zhao Z., Xu K., Hui G. and Hu L., "An EnergyEfficient Clustering Routing-Protocol for WSNs based on AGNES with Balanced-Energy ConsumptionOptimization", MDPI Sensors, 2018, doi:10.3390/s18113938.

[20] Sreedharan, Pete, 2020, "Sp. Clustering Algo. with Dynamic Multi-Hop Routing for WSNs", Int J Emerging Technologies, vol. 11(3), pp. 1169-1178, 2020.

[21] Sreedharan, Pete, "Dynamic Multi-Hop Spatial Clustering Scheme for Data-Transfer in WSN with Enhanced QoS", IEEE Int Conf on Emerging Tech, 2020, doi: 10.1109/INCET49848.2020.9154118.

[22] Qian J, Yu Z, S. Liu and W. Chong, "A Pre-Selected Clustering Protocol Based on Spatial Correlation", 4th International Conference on Wireless Communications, Networking and Mobile Computing, Dalian, pp. 1-4, 2008

[23] Xing, W., Liu, Z., Zeng, B., Wang, Y., Lu, D., "Distributed Spatial Correlation-based Clustering for Approximate Data Collection in WSNs", IEEE 27th International Conference on Advanced Information Networking and Applications, 56-63.

[24] Yin Y, Liu F, Zhou X, Li Q., "An Efficient Data Compression Model Based on Spatial Clustering and PCA in WSNs", *Sensors*. 2015; 15(8):19443-19465.

[25] Rajeev Kumar1, Naveen Chauhan2 and Narottam Chand3, "Adaptive Clustering in Wireless Sensor Networks", International Journal of Multimedia and Ubiquitous Engineering, Vol.12/9, pp.15-26, 2017.

[26] P. Sreedharan, Pete, "Spatial Clustering with Sequential CH Selection for Energy-Efficient WSN", Int. Conf. on Wireless Commun, Springer Lecture Notes on Data Engineering and Communications Technologies 36, pp. 978-981, 2019.

[27] Wang, Liu, Wang, Nie, Bian, "Spatial Query Centric Geographic Routing Protocol in WSN", MDPI Sensors, 2019.

[28] S. Madan. "Clustering of Sensor Nodes using SpatioTemporal Correlation", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5/11, 2017.

[29] Chongqing Z., "Cluster based Routing Algorithms Using Spatial Data Correlation for Wireless Sensor Networks", Jounal of Communications, Vol. 5/3, 2010, doi:10.4304/jcm.5.3.232-238

[30] He, S., Li, Q., Khishe, M. *et al.* The optimization of nodes clustering and multi-hop routing protocol using hierarchical chimp optimization for sustainable energy efficient underwater wireless sensor networks. *Wireless Netw* **30**, 233–252 (2024). https://doi.org/10.1007/s11276-023-03464-9

[31] Rezaii MohammadAli, Azhdari Majid, Mahmoodzadeh Azar, Khishe Mohammad. Power consumption optimization in underwater wireless sensor networks based on EECRU clustering algorithm for routing. Iranian Journal of Marine Science and Technology. 2020;24(95):1–10.

[32] Yang Yang, Yunqiang Wu, Hongji Yuan, Mohammad Khishe, Mokhtar Mohammadi, "Nodes clustering and multi-hop routing protocol optimization using hybrid chimp optimization and hunger games search algorithms for sustainable energy efficient underwater wireless sensor networks", Sustainable Computing: Informatics and Systems, Volume 35, 2022, 100731, ISSN 2210-5379, https://doi.org/10.1016/j.suscom.2022.100731.

[33] Khisheh M., Aghababaei M., Saffari A., Goldani A., "AUV's sensor selection by using ant-lion optimization algorithm and neural networks", Iranian Journal of Marine Science and Technology. 2016; 20(77):59–69.

[34] Khisheh M., Aghababaie M., Saffari A., "Design of AUV sensor selection strategy by using biogeography-based optimization algorithm and neural network", Conference: the Fourth Iranian Conference on Engineering Electromagnetics (ICEEM 2016) 2016.

Houman K, M. G. Nejati, A. Saffari, M. Khishe, M. Mohammadi, "Feature Selection and Training Multilayer Perceptron Neural Networks Using Grasshopper Optimization Algorithm for Design Optimal Classifier of Big Data Sonar", Journal of Sensors, 2022, https://doi.org/10.1155/2022/9620555