Node Localization Using Modified Wild Horse Optimization and Energy Efficient Secured Routing Protocol for IoT Based Wireless Sensor Networks

M. Prakash and Dr.A. Prakash

Abstract--- *Internet of Things (IoT) devices are being utilized extensively as a result of the development of information and communication technology. Wireless sensor networks (WSNs), which are formed of inexpensive smart devices for information collection, plays a vital function in the establishment of the IoT. These smart devices are not without limitations, though, whether it comes to processing, memory, computing, and energy usage. In addition to these limitations, the core difficulties facing WSN include node localization, reliability, and data security during transmission in a dangerous environment from hostile nodes. The important and difficult problem for researchers to solve in order to improve network longevity, reliability, scalability, connectivity, throughput is accurate localization and routing. To improve the network period and data trustworthiness, this study intends to design an Energy-Efficient and Secure Routing protocol (EESR) and node localization centered on modified wild horse optimization (LMWHO) for intrusion avoidance in IoT utilizing WSN. Initially the suggested protocol bases its creation of various energy-efficient clusters on the inherent characteristics of nodes. Secondly, the base station (BS) and cluster head are able to reliably and securely share sensory data according to the (k,n) threshold-based Shamir secret sharing method. And finally, node localization centered on modified wild horse optimization (EESR-LMWHO), where the fitness function was formed by the development of residual energy and distance estimate. The suggested EESR-LMWHO utilizes less energy and prolongs the life of wireless networks. Lastly, the simulations are run to evaluate the suggested method's efficiency. The suggested approach, according to the experiments, approximates the location of the unknown node, offers a minimal localization error, and is a lightweight way to deal with intrusions caused by hostile nodes.*

Keywords--- *IoT, Wireless Sensor Networks (WSN), Node Localization, Secured Routing Protocol, Energy-Efficient and Secure Routing Protocol (EESR), Modified Wild Horse Optimization (MWHO).*

I. INTRODUCTION

The IoT is a global network of communication made up of various connected things that provide networking, sensing, and information processing capabilities. The primary goal of the IoT is to enable interaction among uniform items everywhere, for anything, and at any time [1]. An early example of an Internet of Things technology is radio-frequency identification (RFID), that utilizes wireless networking components to automatically communicate identification data to a reader through electromagnetic fields. The two primary components of an RFID system are tag readers and radio signal transponders (tags). RFID tags typically include electronically stored data that allows users to categorize, follow, and keep an eye on the objects [2]. Any object can have RFID tags affixed to it in order to collect data and track the target location.

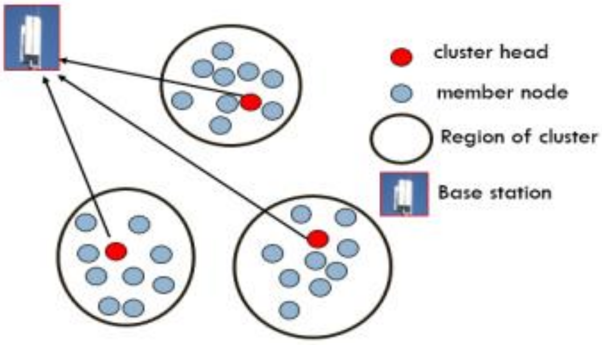

Figure 1: The Structure of the Cluster Based Network

Wireless Sensor Networks (WSNs) are dispersed throughout the environment and comprise a multitude of tiny sensor nodes embedded with sensors, microcontrollers, and batteries on a single chip. These nodes are placed throughout the network's target region to sense and collect data [3]. Because of this, WSNs are highly suitable for applications involving environmental monitoring, such as alerting forest guards to a fire by sending out warning signals. Additional applications include traffic, weather and agricultural monitoring and security surveillance. These days, WSNs are utilized everywhere, including houses, to monitor the temperature of a room and turn on the air conditioner automatically when it exceeds the ideal threshold value [4]. Of course, it might be utilized to track the flow of water and gas, security, making smart houses practical and noteworthy.

Furthermore, IoT-based WSNs are employed in both attended and unattended settings, including smart cities,

M. Prakash, Research Scholar, Department of Computer Science, Hindustan College of Arts & Science, Coimbatore. E-mail: guru8026@gmail.com
Dr.A. Prakash, Professor, Department of Computer Science, Hindustan College of Arts and Science, Coimbatore. E-mail: prakashankar75@gmail.com

DOI: 10.9756/BIJNTA/V11I1/BIJ24004

water quality, and air pollution. In addition to ensuring dependable data transfer, energy efficiency must be increased. A cluster-based approach to WSN energy efficiency has previously been given to several academics [5]. Sensor nodes are powered by tiny batteries that are hard to replenish and have limited processing and communication power. Thus, for these sensor nodes to have a longer network lifespan, their energy consumption should be at least. The energy of sensor nodes is the main constraint when developing WSNs, yet these nodes are also limited with regard to processing capability, storage, transmission range, and energy [6]. Over the past ten years, it has been discovered that cluster-based routing is a technique that can effectively decrease sensor nodes' energy use and increase network lifespan when equated to other approaches like direct transmission. Comparing with other methods, clustering offers a network lifespan that is two or three times longer. Sensor nodes are grouped together to form clusters during clustering, which reduces the amount of long-distance sensor node transmissions and saves energy [7]. Further, this leads to energy savings as CH inside each cluster has responsibility for each cluster member sensor node [8]. By limiting the amount of data delivered, data aggregation at CH also contributes to sensor node energy savings.

Node localization technology is able to find and track nodes, making the monitoring data more relevant. In this case, without the localization data of the nodes in the sensor field, the user would not be able to understand the data collected at the sink node [9]. Localization is the process of determining the position of unknown sensor nodes, employing the known position of sensor nodes, utilizing measurements such as arrival time, arrival time difference, arrival angle, maximal likelihood, triangulation. Global positioning systems (GPS) with every sensor node can solve the localization problem of WSN, although this is not recommended because of energy, cost, and size concerns [10]. It even performs poorly underwater and indoors. Thus, a more effective and superior solution is required to localize the sensor nodes. There are numerous non-GPS localization methods that constitute two groups: range-based and range-free methods. Point-to-point or angle-based calculation of distance among sensor nodes is employed by range-based localization techniques [11]. This utilizes trilateration of anchor nodes to estimate location. Range-free localization methods rely on topological information instead of range information amongst the anchor and target nodes. Range-based localization methods are less cost-effective than range-free methods, but they offer greater precision. This work intended to establish an EESR and node localization centered on modified wild horse optimization (LMWHO) for WSN-based IoT intrusion prevention to lengthen network lifespan and improve data reliability.

The following study is arranged as follows: section 2 examines node localization and energy-efficient routing strategies. The suggested approach is presented in section 3. The findings and discussion are given in section 4. Section 5 discusses the conclusion and further research.

II. Literature Review

This section reviews the some of the recent energy efficient routing and node localization methods in WSN.

Shi et al [12] proposed a novel secure routing scheme for WSNs when hostile nodes are available. The protocol considers associated data. The trust value is definite as the node's attack probability determined by prior packet-forwarding activities, and the status is a hybrid measure that incorporates the remaining energy and distance. The linked data indicates that the route that the protocol generates is globally optimum and secure against malicious attacks. Employed an enhanced Dijkstra algorithm version to produce the route that WSNs should take when malevolent nodes are accessible. The suggested paradigm's ability to sustain a higher delivery ratio when compared to the RBMSC algorithm in the same simulated setting confirms its effectiveness based on global optimization. Khan et al [13] presented a practical, trustworthy routing technique built on the hybrid trust framework to combat self-centered nodes. The TASRP is a multifactor routing technique which employs residual energy, path length, and node trust scores to produce dependable routing paths among trusted nodes while using less energy. Because its routing paths are shorter, this multi-factor technique helps choose reliable nodes to forward data and save energy usage. Improved efficiency is demonstrated by the simulation findings of node energy consumption, throughput, and robust trust values. Selvi et al [14] suggested a secure routing method, that utilizes spatiotemporal constraints in conjunction with a decision tree process to establish the optimal route and effectively utilizes trust score evaluation to identify rogue users in wireless sensor networks. The findings show that in terms of security, and packet delivery ratio, the optional approach performs better than the modern techniques.

Kalidoss et al [15] suggested an Energy Efficient Routing Protocol with QoS monitoring that is built on trust and energy modeling to maximize energy usage while bolstering WSN security. Trust modeling combines an authentication mechanism with a key-based security technique that generates trust scores. In order to improve communication security, a trio of kinds of trust scores are computed. A cluster-based secure routing method is suggested, whereby the cluster head is chosen to execute cluster-based secure routing depending on trust scores and QoS parameters. To ensure the effective execution of the safe routing procedure, the final path was taking into consideration path-trust, energy, and hop count. Liu et al [16] developed a trust routing and security system for WSNs called ActiveTrust. The main advancement of ActiveTrust is in its ability to prevent black holes by actively generating several detection routes, which enable prompt nodal trust acquisition and enhance data route security. The generation and dissemination of detection routes is rendered possible by the ActiveTrust system, enabling the full utilization of the energy found in non-hotspots to generate multiple detection routes as necessary to achieve the necessary levels of security and energy conservation. Sivasakthiselvan et al [17] suggested the LMS coefficients scheme in a modified evolutionary model is employed to improve MCL. To select the best sample sets and drastically decrease the number of communication hops, which lowers network traffic, communication overhead, and localization error. The findings indicate that the approach achieves better than the Monte-Carlo Localization and the Kalman Filter Localization technique.

Kulkarni et al [18] developed an iterative, distributed localization. The nodes that are localized during an iteration serve as guides for the localization of the remaining nodes. PSO and BFA have been employed to solve the issue. The quantity of nodes localized, localization accuracy, and computation time are compared between PSO and BFA's characteristics. Arora et al [19] introduced A butterfly optimization technique that suggest a node localization method. The suggested method is tested by simulating sensor networks with varying densities, from 25 to 150 nodes, whose distance readings are tainted by gaussian noise. The suggested innovative scheme's efficiency is contrasted with that of a few recognized strategies, including FA and the PSO algorithm. According to the outcomes, the suggested strategy outperforms the current PSO- and FA-based node localization strategies in terms of consistency and accuracy, node location. Kanoosh et al [20] proposed a node localization technique centered the recently developed SSA bioinspired technique. Over various WSN deployments, the suggested approach is contrasted with recognized optimization methods. According to the findings, the suggested method outperforms the alternative approaches in regards to the quantity of localized nodes, and computation time.

Cheng et al [12] suggested an efficient node localization CS method. This method utilizes the fitness of each solution to create the mutation probability, and it allows the population to rapidly approaches the global optimal solution according to the variation of step size. Additionally, the system restricts the population within a specific range to avoid the energy usage resulting from pointless searches. Numerous tests were run in order to determine how the suggested method would perform in localization success ratio and average localization error were changed. A comparison analysis was carried out to accomplish the identical localization task with the similar network deployment. The outcomes demonstrate that the suggested CS method, when contrasted with the traditional CS and PSO method, can decrease average localization error in addition to increasing convergence rate. Goyal et al [22] defined the bat algorithm is used to assess the accuracy of the node localization issue in WSN. In the meantime, the bacterial foraging techniques of the bacterial foraging optimization method were further utilized to modify the current bat algorithm. Simulations demonstrate that, when compared to the current method, the suggested system consistently works better according to improving resilience as well as localization success rates and quick convergence times.

III. Proposed Methodology

This section presents an outline of the modified wild horse optimization (LMWHO)-based node localization and EESR protocol for ToT-based WSN. Figure 1 depicts the fundamental layout of a cluster-based network. Three key elements comprise the general functioning of the suggested EESR procedures, which are examined here.

1. The thresholding-based secret sharing scheme (SSS) and optimal hierarchical topology building are arranged by the EESR protocol in the initial element to provide safe data routing. The optimized cluster heads are identified in relation to the distributed clusters in a balanced and energy-efficient manner utilizing a number of factors and QoS restrictions. Additionally, the suggested clustering approach increases the power utilization ratio and network lifespan with minimal overhead among the sensor nodes.
2. To prevent intrusions induced by hostile nodes, a reliable and secure routing path is built among cluster heads and the BS in the second element. Selected cluster heads share a secret key that the BS produces to ensure dependable data transmission. The SSS technique is employed to encrypt data packets during data passing from cluster heads. However, BS utilizes the suggested secret sharing mechanism to reconstruct the incoming data packets from cluster heads.
3. And finally, node localization centered on modified wild horse optimization (EESR-LMWHO), where the fitness function is formed by the distance estimation and residual energy.

The general method of the suggested methodology is illustrated in figure 2.

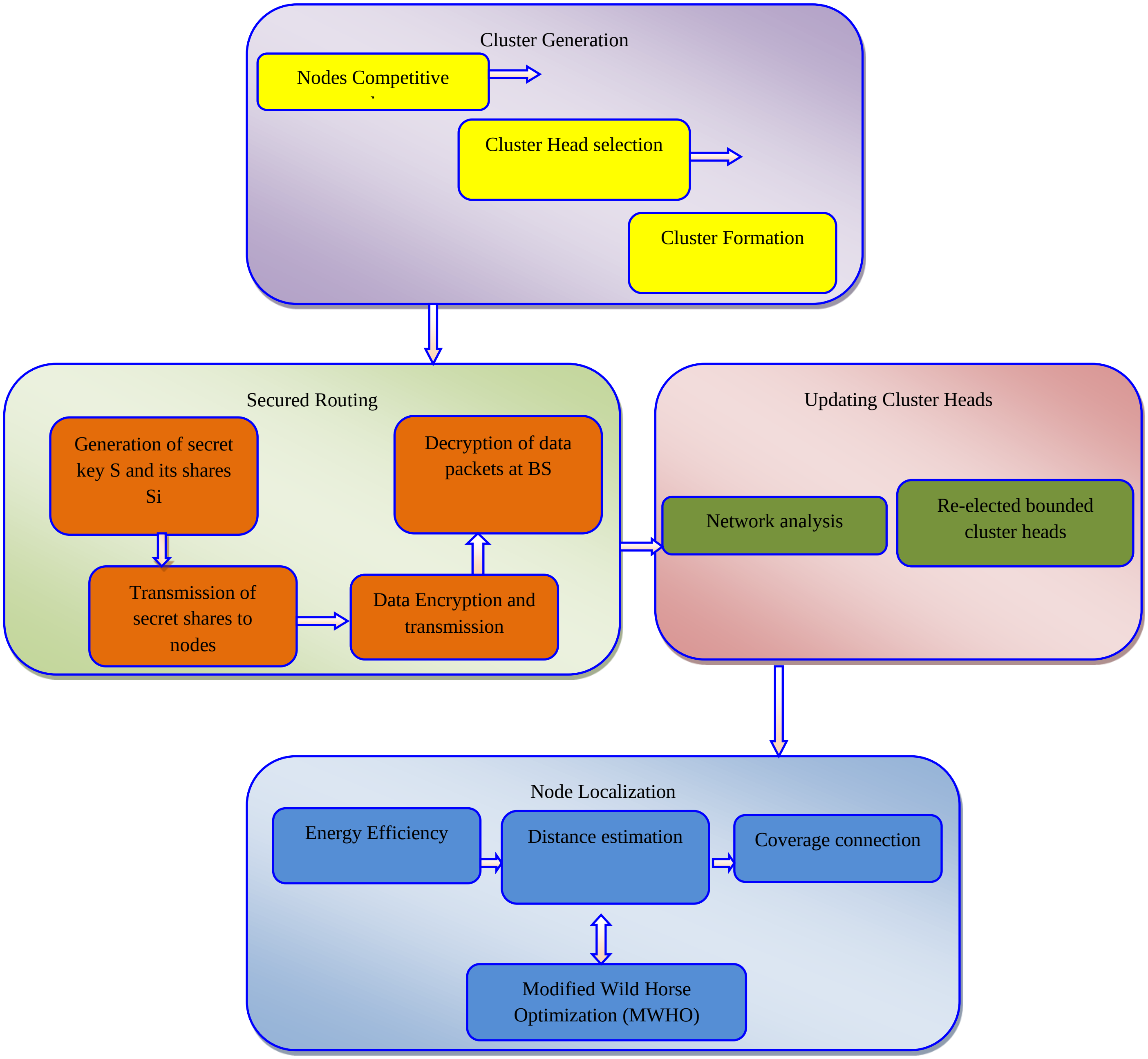
Cluster Generation
Nodes Competitive
Cluster Head selection
Cluster Formation
Secured Routing
Generation of secret key S and its shares Si
Decryption of data packets at BS
Transmission of secret shares to nodes
Data Encryption and transmission
Updating Cluster Heads
Network analysis
Re-elected bounded cluster heads
Node Localization
Energy Efficiency
Distance estimation
Coverage connection
Modified Wild Horse Optimization (MWHO)

Figure 2: Block Diagram of the Proposed Methodology

Optimal Clusters

In the startup step, the number of nodes is distributed at random across a square-sized network field. Every node has a unique ID with few restrictions and stays stationary. The boundless resources are not restricted for the BS. Initially, BS uses the monitoring field to broadcast its location hop-by-hop, and every node receives it. Additionally, the information from the neighbor is added to each node's routing table. Subsequently, the ESR protocol disperses its announcement of the cluster head selection procedure. The competitive value C_v is calculated for each node utilizing residual energy (ei), the Received Signal Strength Indicator $(RSSI_i)$, the queue length QLi_i, and the proximity from BS di to BS. Every node exchange control message with its neighbor to obtain information about each other. Firstly, node energy plays a major role in the network's survival; as a result, a node's maximum residual energy is given more weight. Secondly, if the RSSI value is higher than a predetermined threshold, the RSSI is employed to assess the efficiency of the wireless link that provides a decent packet reception rate. Equation 1 provides beacon packets average reception rate from N neighbors at a specific time period (Δt), which is the RSSI threshold that the ESR protocol estimates. The determined threshold needs to be exceeded by the node's RSSI value. Low connection quality is indicated by an RSSI value below the threshold, which raises the probability of a packet loss ratio. Let X represent the beacon packet reception rate. Next, the RSSI threshold can be determined with the following formula.

$$RSSI_{threshold} = \frac{X}{N} \qquad (1)$$

Thirdly, the network lifetime and energy usage are extended when a node takes the shortest way to the base station. Ultimately, the queue length QLi_ifactor enhances the efficiency of data delivery and gauges the degree of congestion at the node level. Equation (2) might be utilized to calculate the queue length $QL_i$of a node i where RR_i is the number of packets received in bytes at node i and TB is the total buffer size in bytes.

$$QL_i = \frac{RR_i}{TB} \qquad (2)$$

Equation (3) displays the weighted averages of all the factors, which are then summed. The nodes are then designated as the initial cluster heads according to the maximum competitive value Cv. As a result, the suggested ESR process chooses an optimal cluster head considering inherent characteristics, and the resulting clusters are more flexible. The calculated Cv value is normalized within the interval [0,1].

$$C_v = w_1 \times e_i + w_2 \times RSSI_i + w_3 \times \left(\frac{1}{d_{i\ to\ BS}}\right) + w_4 \times QL_i \ (3)$$

Weighting considerations for various selection aspects (i.e., the node's residual energy, RSSI, proximity from BS, and queue length) are represented by w_1, w_2, w_3 and w_4 in Equation (3). In the method of selection, each weighting element represents a specific influence on determining the competitive worth of nodes, while $w_1 + w_2 + w_3 + w_4 = 1$. Since the values of the queue length, RSSI, residual energy, proximity from the base station, and RSSI are all within the same range, the estimated competitive value is in the range of [0,1]. The cluster selection technique is rendered more adaptive by the residual energy metric. The cluster head's selection system, which displays the wireless link's efficiency, integrates the RSSI facet. An appropriate node is measured for the cluster head selection according to the shortest distance from the base station. Every node sends out beacon packets to its neighbors on a predetermined interval. The neighbor node analyzes the beacon packets it receives, determines their RSSI value, and then sends them back to the source node. Ultimately, the cluster head selection procedure takes queue length into account; if a node's transit queue length exceeds a predetermined threshold, it will be given a greater priority to designated as a cluster head. Following the primary cluster chiefs' selection, they made a precise announcement about their status. After receiving the status updates, all of the regular nodes join their nearby cluster head to form clusters. Regular nodes may associate with the cluster heads that have the highest RSSI value after receiving status messages from multiple neighboring cluster heads. Upon completion of the cluster formation procedure, each produced cluster is given a unique ID by the ESR protocol, which helps to define its borders. Time-division multiple access (TDMA) centred channel access schedules are announced by the group of nodes chosen as cluster heads.

(t, n) Thresholding Based Secret Sharing Scheme (SSS) for Secure Data Routing Against Intrusions of Malicious Nodes

The suggested method utilizes a (t, n) threshold centred Shamir's secret sharing system where a secret key S is generated by the BS and divided across a group of n cluster heads. Any t subset of cluster heads is sufficient to recreate the secret key S. It should be mentioned that two requirements need to be met in order for Shamir's secret sharing plan to be implemented:

i. Any set of t or more subkeys $(S_0, S_1, \dots, S_{t-1})$ might be utilized to rebuild the secret key S.

ii. The secret key S cannot be reconstructed using fewer than t subkeys. In SSS, a t − 1-degree polynomial construct t subkeys. A (t, n) threshold system is constructed by choosing t − 1 random numbers $(b_0, b_1, \dots, b_{t-1})$ that are greater than zero. The values $(b_1, b_2, \dots, b_{t-1})$ are the polynomial's coefficients if $b_0 = s$, as stated in Equation (4).

$$f(x) = b_0 + b_1x + b_2x^2 + b_3x^3 +, \dots\dots, b_tx^{t-1} \qquad (4)$$

The Lagrange basis polynomial stated in Equation (5) must be computed to rebuild the secret keys S.

$$l_j(x) = \prod_{\substack{0\le m\le t\\ m\ne j}} \frac{x-x_m}{x_j-x_m} \qquad (5)$$

Following the computation of t-1 Lagrange values, entered into equation (6) to determine secret key S.

$$f(x) = \sum_{j=0}^{t-1} y_j l_j(x) \qquad (6)$$

Every portion of the key S_i is dispersed among the cluster heads and then flooded toward a single cluster node. The sensory data D_i is protected by executing the Exclusive OR (XOR) operation with the key S_i, as stated in Equation (7), when the node communicates it to the cluster head.

$$E_i = S_i \oplus D_i \qquad (7)$$

The cluster head sends the encrypted data E_i to the BS for additional processing after obtaining it from the member nodes. According to this, when encrypted data arrives, the BS uses the decryption key S to decrypt it before sending it to the user.

Updating of Cluster Heads

The ESR protocol dynamically recasts the role of cluster chiefs because WSNs have limited resources. Ensuring consistent load balancing and energy usage is the primary goal of updating the cluster heads element. The following is observed by the EESR protocol to assess the network metric.

i. When a data packet is received, each cluster head j checks to see if it previously acquired the similar packet or not. If so, cluster heads just discard the redundant data packet to save energy and network congestion.

ii. When the cluster head gets a novel data packet but is unable to process it due to insufficient energy (e.g., ej < threshold), it may stop data forwarding and start the re-election mechanism inside a specific cluster limit. In addition, the congestion rate Cr of every cluster is calculated by the EESR protocol utilizing the function, which represents the normalized congestion value in the interval [0,1] as indicated by Equation (8).

$$C_r = \frac{ADR}{ARR} \qquad (8)$$

The average reception rate of data packets is represented by ARR, and the average delay ratio is denoted by ADR. The C_r value is continuously checked by the ESR protocol. If it does not fall within the range of [0,1], the protocol considers that a certain cluster head has exceeded the congestion limit and that the re-election procedure needs to be started.

Node Localization

The node localization centred on modified wild horse optimization (LMWHO), where the fitness function is formed by the distance calculation and residual energy.

A. Description of Localization

Received signal strength (RSS) measures are commonly employed in real-world localization apps and are authorized as a conserved restricted measurement in the majority of studies due to their inexpensive cost and ease of hardware implementation. The following is a statement of the RSS at a trustworthy location inside a localization area:

$$RSS(d_0) = P_t + K_e - 10\eta \log\left(\frac{d_0}{d_1}\right) + \alpha + \beta \qquad (9)$$

where P_t represents the nominal transmission power (dBM), K_e represents a constant based on the system, η means the path loss coefficient, d_1 is a far-field antenna's reference distance, α signifies the fast fading effect, and β signifies the random reduction resulting from shadowing. By modifying d_0, the actual transmitter-receiver distance, the RSS is examined. Anchor nodes are those whose exact locations are known before the localization process begins. Alternative names for these are known nodes. A collection of all WSN nodes with known locations is denoted as *KN* when *A* is the number of anchor nodes. Thus, (p_{kx}, q_{kx}) represents a known node position K_x. Additionally, unknown nodes are those whose position is determined by a specific localization technique. *UN* is employed to indicate the group of unidentified WSN nodes.

$$\left.\begin{array}{r} KN = K_x | x = 1,2, \ldots\ldots, A \\ UN = U_x | x = 1,2, \ldots\ldots, B - A \\ RN = E_x | x = 1,2, \ldots\ldots, C \end{array}\right\} \qquad (10)$$

Let B − A be the number of unrecognized nodes. The true positions of U_x described by (p_{ux}, q_{ux}) are undetectable in a real-time request. Assuming that the communication range has a radius of R. Since p_x and p_y are two examples of powered sensor nodes, p_x is immediately regarded as p_y's neighbor if it is situated within p_y's broadcasting range. Thus, p_y is a neighbor of p_x as long as all the activated WSN nodes have relatively comparable transmission ranges. Since other unknown nodes are identified throughout the placement, the probability of finding a node without a precise location is provided as an estimate. The estimated position of U_x is signified by (p_{ux}^0, q_{ux}^0). Localization seems to be done with the intention of $(p_{ux}^0, q_{ux}^0) = (p_{ux}, q_{ux})$ for U_x. The reference nodes are obtained from localized unknown and known sensor nodes in the process of trying to locate an energized sensor node. The array of reference nodes is signified as *RN*, where *B* ≥*C* ≥*A*. The reference node point E_x with an exact position is definite as (p_{ex}, q_{ex}). If *Ex* is anchoring K_y, then $(p_{ex}, q_{ex}) = (p_{ky}, q_{ky})$. Yet, if E_x comprises localized unknown nodes U_k, at the moment $(p_{ex}, q_{ex}) = (p_{uk}^0, q_{uk}^0)$.

The actual distance d_{xy} is the distance travelled among the real locations of U_x and E_x. utilizing the error z, which is established by the random employed measuring instrument, a specific measurement procedure yields the measurement distance d_{xy}^1. This error is typically replaced by random value in subsequent research for convenience. Here presuppose that $d_{xy}^1 = d_{xy} + N(0, d_{xy}Z)$, where $(N(0, d_{xy}Z))$ is the Gaussian function with a mean value of 0 and a variance of $d_{xy}Z$. The distance enclosed by the assessed positions U_x and E_y is signified as d_{xy}^0. Assume unknown node U_x has m neighbor reference nodes $E_1, E_2, \ldots, E_m$, here $y = 1, 2, \ldots, m$. To obtain (p_{ux}^0, q_{ux}^0), derive the subsequent equations:

$$d_{xy}^1 = \sqrt{(P - P_{ey})^2 + (q - q_{ey})^2} \qquad (11)$$

where (*p*, *q*) is an unknown dimension to solve and P_{ey}, q_{ey} is the position of E_y. According to the presence of the estimated position (p_{ux}^0, q_{ux}^0) and the distance measurement error *z*, it is not feasible to ascertain the precise location of E_y in a general sense. Using the estimated position (p_{ux}^0, q_{ux}^0), express the estimated distance d_{xy}^0 at that location as

$$d_{xy}^0 = \sqrt{(p_{ux}^0 - P_{ey})^2 + (q_{ux}^0 - q_{ey})^2} \qquad (12)$$

Since the actual distance varies due to the uncertainty of *dxy*, positioning attempts to get the smallest possible distance between $d_{xy}^1 to\, d_{xy}^0$. Lastly, create the location issue represented by U_x as

$$\sum_{y=1}^{n} w_y \,(d_{xy}^0 - d_{xy}^1)^2, \qquad (13)$$

$$\sum_{y=1}^{n} w_y \left(\sqrt{\left(P - P_{ey}\right)^2 + \left(q - q_{ey}\right)^2} - d_{xy}^1\right)^2 \qquad (14)$$

where $w_y = (1/d_{xy}^1) \sum_{x=1}^{m} (1/d_{xy}^1)$,which provides a clearer grasp of the reference point nearer U_x. In actuality, asymmetrical message delivery, multipath fading, and redundant noise generate the midway circle that encircles the energized sensor communication range. The variance among the calculated and real location points must considered when determining the lowest localization error of unidentifiable location points U_x,

$$LE_x = \frac{1}{R}\sqrt{(p_{ux}^0 - P_{ux})^2 + (q_{ux}^0 - q_{ux})^2} \qquad (15)$$

B. Clustering Model

Here, determines the fitness function for the suggested node localization using a modified wild horse optimization algorithm.

The closest point among the energized sensors and the closest transmission range are found to be in the same locality (cluster) through the node clustering process, which is meant to preserve energy. The recommended clustering technique is put up in Figure 3. Finding the exact position is the ,e key concern, and finding it requires making a number of judgments. The sensor node distance is computed utilizing (5) to establish the ideal position for a specific energized sensor. The innovative method of structured clustering divides the whole WSN nodes into multiple clusters by linking sensor nodes with Euclidean distance.

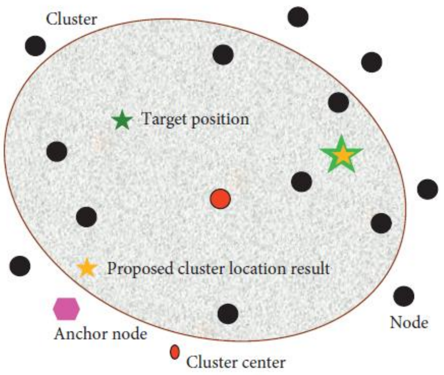

Figure 3: Expected WSN Cluster Structure

- *Energy Efficiency*

A horse's resilience to all circumstances contributes in reducing energy loss and extending the life of the network. The whole distance among the receiver and transmitter still taken consideration in the channel description for multipath fading and free-space. A multipath decay model is employed if d is more than or equal to d_0. If the upper threshold value d_0 is less than the node distance pairs d, the energy amplification consumption is assumed to follow a free space model. As a result, the radio's energy need to transmit an h-bit message over a distance of d is provided in (16). ,e radio utilizes up energy to accept a h-bit message given in (17). E_{elec} reliance on elements like signal dispersion combining, digital coding, filtering, and modulation, but the energy required to magnify the system, $\varepsilon_{fs}d^2$ or $\varepsilon_{mp}d^4$, depends on the appropriate error per bit and the reception structure according to the distance traveled. The electronic energy utilized by the electronic circuit is known as E_{elec} and $d_0 = \sqrt{\varepsilon_{fs}/\varepsilon_{mp}.\varepsilon_{fs}}$ are the amplifier energies in multipath and free space, individually. When a message is transmitted over hth-bits at a distance d from the recipient in equation (18), define E_i as the remaining energy. E is the recent energy of the node. The energy required to transmit a message plus the energy utilized to receive a message is then $E_1 = E_T(h,d) + E_R(h)$.

$$E_T(h,d) = \begin{cases} h \times E_{elec} + h \times \varepsilon_{fs}d^2, & d < d_0 \\ h \times E_{elec} + h \times \varepsilon_{mp}d^4, & d \geq d_0 \end{cases} \tag{16}$$

$$E_R(h) = h \times E_{elec} \tag{17}$$

$$E_i = E - E_1 \tag{18}$$

- *Distance Estimation*

The entire distance traveled begins at an activated node point and ends at a different sensor point, that is assumed to be the separation between both adjacent sensor nodes. The expression for it is $dist(p_a, q_b)$. The distance between a locational node and the central cluster position need to be negligible in comparison to the distance between the cluster center and another node. The former is written as (p_c, q_b), and the latter is formatted as $dist(p_a, q_c)$. This increases the cluster strength and decreases the lack of sensor node engagement; thus, it must be ensured in an organized manner to minimize the energy waste of each node in a large network. Calculate the sum of the distances D_i and q_y between each node point $p_x \varepsilon N$, here N is the collection of all nodes. Nevertheless, these energetic surrounding spots might represent a locationless node that is activated by a known node's position. (21) provides this distance.

$$D_G^0 = \sum_{p_a \in N}^{N} dist(p_a, q_b), \tag{19}$$

$$D_G^1 = \sum_{q_b \in N}^{N} dist(p_a, q_c) + dist(p_c, q_b) \tag{20}$$

$$D_i = \sum_{p_y \in N} dist(p_x, q_y) \tag{21}$$

- *Coverage Connection*

Each WSN may be viewed as a connected undirected figure, represented by $G = (V, E)$, where V is composed up of vertices that include the E edge set $\{e_1, e_2, \dots \dots . . e_f\}$, which represents the distance among the energized sensor nodes, and $\{v_1, v_2, \dots \dots . . v_u\}$ which represents the energized node point located in the WSN. This method requires the weighted values that are dependent on coverage connection, distance estimate, and energy efficiency and are shown on the edges. Each edge within the network has a limited real number, denoted by w_i. Let S_r represent a node's sensing range. The connection variables $c = c_1, c_2, \dots \dots . . c_m$ should be identified with the activated sensor nodes $p_x \, and \, p_y$. But C is the WSN area, N is the total of all identified energized sensor nodes, and C_y is the area represented by the yth cluster center node.

$$C_L = \begin{cases} 1, & if \left\| p_x - p_y \right\| \leq S_r \\ 0, & otherwise \end{cases} \tag{22}$$

$$C_i = \bigcup_{y=1}^{N} \in C_y \frac{C_y}{C} \tag{23}$$

The following represents the final fitness function calculated during minimization; it consists of the earlier fitness minor objectives:

$$F_i = w_1 E_i + w_2 D_i + w_3 C_i \tag{24}$$

where $w = \{w_1, w_2, \dots \dots . . w_f\}$ is the distance linked with the edges. To determine the fitness function's value to each of the other subfunctions, consider w_1, w_2, w_3 as the weight coefficients linked to the function, which are denoted by $\sum_{i=1}^{3} w_i \geq 0, w_i \in (0, 1)$.

C. Modified Wild Horse Optimization (MWHO)

The WHO adopts a wild horse behavior. Non-terrestrial horses are referred to as wild horses. They reside in two groups: a family group for mares, or female horses, and a separate group for stallions, or male horses. Among the family group and the single group, mating occurs. Foals, or young horses, are concerned with grazing when they are first born [23]. After leaving their home group, female foals join other groups. Once a male colt reaches maturity, they are referred to as stallions. Stallions, yet, link the "single group" out of civility. Decency, which incest is avoided by assembling the stallions into a single group. The habits of dominant leaders, which have access to water holes while lower members have to wait for hours, was proven by their search for water throughout dry seasons. Family groups are led by mares, but as subordinates, they have to submit to a leader chosen by the stallions. The following are the WHO's primary stages:

- *Population Initialization and Leadership Selection*

The initial population$(\vec{x})$ consisting of N individuals is randomized in such a way that $(\vec{x}) = \{\vec{x}_1, \vec{x}_2, \dots\dots\dots \vec{x}_n\}$. Each population's objective function is then computed to create its associated vector.

$$\left(\vec{O}\right) = \{\vec{O}_1, \vec{O}_2, \dots\dots\dots \vec{O}_n\} \qquad (25)$$

Utilizing PS to represent the stallions percentage in the total population, the population is separated into groups G = NXPS. At the beginning of the method, each group has a randomly chosen stallions leader; however, as the method progresses, the highest fitness value determines which leaders are chosen.

- *Grazing Behavior*

In Equation 26, the grazing habit is depicted.

$$\bar{X}_{i,G}^{j} = 2Zcos(2\pi RZ) \times \left(Stallion^{j} - X_{i,G}^{j}\right) + Stallion^{j} \quad (26)$$

where $X_{i,G}^{j}$ is member's group current position, $Stallion^{j}$ specifies the group leader's position, the Z variable is supplied by Equation (28), R is a uniformly distributed random number in the interval [−2, 2], and the horses graze around the group leader at diverse angles (360 degrees). π is taken as 3.14, the movement in distinct radii is caused by the cosine function of R and π, and the updated position of a member $\bar{X}_{i,G}^{j}$ is its final position.

$$P = \vec{R}_1 < TDR: \qquad IDX = (P == 0); \quad (27)$$

$$Z = R_2 \ominus IDX + \vec{R}_3 \ominus (\sim IDX) \qquad (28)$$

here P is a vector ∈ [0, 1], $\vec{R}_1$ and $\vec{R}_3$ are random vectors ∈ [0, 1], R_2 states a random number ∈ [0, 1], the $\vec{R}_1$ yields' IDX indexes which fulfill the standards (P == 0). TDR reduce from 1 to 0, as shown by Equation 29.

$$TDR = 1 - iter \times \left(\frac{1}{maxiter}\right) \qquad (29)$$

- *Horse Mating Behavior*

Equation 30,31,32 describes decency and mating behavior.

$$X_{G,K}^{P} = Crossover(X_{G,i}^{q}, X_{G,j}^{z}) \qquad (30)$$

$$i \neq j \neq k \; p = q = end, \qquad (31)$$

$$Crossover \; = \; Mean \qquad (32)$$

where $X_{G,K}^{P}$ denotes horse p's place as it departs group k is replaced by a horse whose parents depart groups i and j as a result of puberty. They have mated and produced $X_{G,i}^{q}$, yet they are unrelated to one another. When the foal q reached adolescence, it mated with the horse z in position $X_{G,j}^{z}$, leaving the j group. The foal q is in the i group.

D. Group Leadership

The group under the leader must guide the others to the proper area of the water. This water is contested over by leaders for its usage of the dominating group; other members are not allowed to utilize until the dominating group has left. Equation 32 illustrates this trend in the same way as (33), where WH is the water position, Stallion Gi is the group i's present leader's position,

$$\overline{Stallion}_{G_i} = \left\{\begin{matrix} 2Ccos(2\pi RZ) \times \left(WH - Stallion_{G_i}\right) + WH & if \; R_3 > 0.5 \\ 2Ccos(2\pi RZ) \times \left(WH - Stallion_{G_i}\right) - WH & if \; R_3 \leq 0.5 \end{matrix}\right\} (33)$$

- *Exchange and Leadership Selection*

At the start, the leaders are chosen at random. At a later step of the process, the population that is the fittest is chosen to be the leader. The positions of the chosen member and the leader are displayed in (34),

$$Stallion_{G_i} = \left\{\begin{matrix} X_{G,i} & if \cos \, t\left(X_{G,i}\right) < \cos t \; (Stallion_{G_i}) \\ Stallion_{G_i} & if \cos \, t\left(X_{G,i}\right) > \cos t \; (Stallion_{G_i}) \end{matrix}\right\} (34)$$

Two techniques are included to the original method to improve its optimizing capability. First of all, wild horses typically rush and pursue for prey. Consequently, the stallions and foals are subjected to a probability random running. The waterhole is then equipped with a dynamic inertia weight, which helps to balance exploitation and exploration.

- *Probability Random Running (PRR)*

Wild horses love to run and chase one another in the wild. This leads to the proposal of the random running approach for stallions and foals equally. The following is the presentation of the position-updating formula:

$$X_{G,j}^{i} or \; Stallion_{G,j} = lb + (ub - lb) \times rand \qquad (35)$$

here *ub and lb* are the upper and lower boundary. During the RRS, search agents could show up anywhere in the search area. So, search agents can break out of the local optima with the support of this technique. It should be noted that the probability of random running (*PRR*) has been given a small value of 0.1 in order to balance exploration and exploitation.

- *Dynamic Inertia Weight (DIW)*

Many studies employ the dynamic weight technique, which helps systems identify the best possible global solution. Therefore, in the first formula of Equation (36), to support the stallions locate an enhanced waterhole, the current inertia weight is increased by a dynamic one. Here's how the weight and adjusted formula are determined:

$$w = \begin{matrix} w \; min + (w \; max - w \; \mathrm{min}) \times \frac{\mathrm{f(t)i - f(t)min}}{f(t)avg - f(t)min}, & \mathrm{if \; f(t)i \leq f(t)avg} \\ w \; max, & \mathrm{if \; f(t)i > f(t)avg} \end{matrix} \quad (36)$$

$$Stallion_{G,j} = 2Zcos2\pi RZ \times WH - Stallion_{G,j} + w \times WH \qquad (37)$$

where w_{min} and w_{max} are the upper and lower boundary values, the present stallion's fitness value at the *t*th iteration is denoted by $f(t)_i$, the average fitness value of all stallions is represented by $f(t)_{avg}$, and the population's minimal fitness value is indicated by $f(t)_{min}$.

The pseducode of WHO is represented in algorithm 1.

```
Algorithm 1: WHO
Input : Raw data
Output: optimized features
Initialization: initialize the parameters PC, PS.
Initialize populations.
Compute the each population’s fitness value
Form groups and select leaders.
while (iter <= maxiter) do
    compute TDR utilizing Equation 28.
    for each stallion do
        compute Z utilizing Equation 27.
        for each foal inside the group do
            if rand > PC then
                update position by Equation 26
            else
                update position by Equation 30
            end if
        end for
        if rand > 0.5 then
            update position of $\overline{Stall\imath on\ G_\iota}$ by Equation 33 first part
        else
            update position of $\overline{Stall\imath on\ G_\iota}$ by Equation 33 second part part
        end if
        if fitness($\overline{Stall\imath on\ G_\iota}$ ) >fitness(Stallion) then
            Stallion = $\overline{Stall\imath on\ G_\iota}$
        end if
        Sort group foals based on fitness levels
        choose the foal with minimum fitness
        if fitness(foal)<fitness(Stallion) then // PRR
            exchange foal and stallion place based on eq 35
            If the stallion's candidate standing is superior
            Applying the potential position from equation 37, change the stallion's location. // DIW
            Exchange foals and stallions position using equation
            t=t+1
        end if
    end for
end while
Return the solution with best fitness
```

Foals and stallions can adjust their positions more easily in MWHO. Search agents may enhance exploration and exploitation with the assistance of the PRR. DIW additionally enables it possible for the approach to provide high-quality solutions and accelerate convergence.

IV. Results and Discussion

Here, NS2 was utilized as a network simulator for the execution and assessment stages. The simulation settings utilised in this investigation are displayed in Table 1. When simulating the proposed EESR-LWHO model, 200 nodes are placed in a 100 m 100 m region. Randomly scattered atypical nodes indicate a lack of interest in data transmission. The sensor nodes have a starting energy of 0.1 Joules. A random mobility model is used to provide sensor nodes motion. Here, the sensor node travels from a certain starting point to the final destination. BS does not move and stays in place. Based on performance indicators including packet delivery rate, average delay, power usage, and network life, the effectiveness of proposed system is evaluated.

Table 1: Simulation Parameters

Simulation parameters	Values
Lengths of Packets	4496 bits
Used energies in transmissions and receipts of Data	50 nJ/bit
Used energies for collecting data	$5*10^{-9}$J
Pause times in interval	0.01 s

PDR Analysis

The basic goal of routing is to guarantee uninterrupted data packet delivery from the point of origin to the final destination. There are instances, nevertheless, in which the sensor node may not be eager to transfer the packet, leading to a decreased delivery rate. It is absolutely normal for a node to choose not to forward any incoming packets when the pace at which it delivers packets is low. The outcomes are displayed in Figure 4.

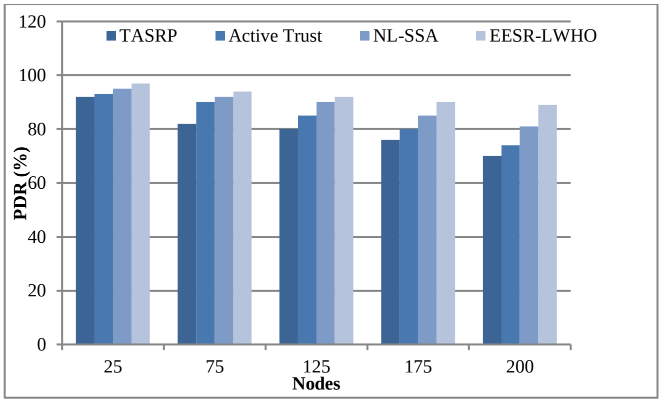

Figure 4: Packet Delivery Ratio

In Figure 4, the packet delivery ratio is displayed. This is because the ESR protocol prioritized the nodes that were better suited to create stable and energy-efficient clusters, taking into consideration a variety of optimal characteristics for choosing the cluster head. Additionally, the (k, n) threshold centered Shamir secret sharing technique was employed to ensure the trustworthiness of the data packets which sent among the BS and cluster head. This reduces route breakages and consequently improves data delivery efficiency.

Average Latency Rate Analysis

Sending a packet to its destination takes time, which is known as latency. Every routing approach needs to have a low latency to increase routing speed. Figure 5 displays the outcomes of the latency tests.

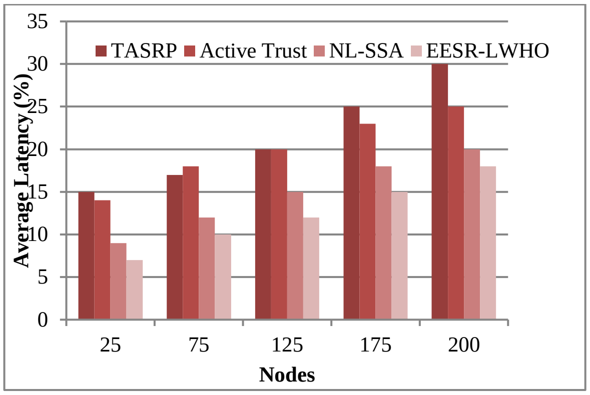

Figure 5: Average Latency Rate Analysis

As shown in figure.5. Given that the EESR-LMWHO protocol specified a lightweight secret sharing structure among BS and cluster heads that was both extensible across a range of different network topologies and computationally secure for managing intrusions towards malicious nodes, according to the computations, the protocol reduced network overhead by an average of 28% when compared to previous solutions.

Energy Consumption

Reducing the routing method's energy usage is essential to extending the network's lifespan. All nodes start with the same amount of energy, and the tasks assigned to them determine how quickly that energy is depleted. The seconds consumed to perform simulations are factored into energy usage computations. The findings for the quantity of energy used are shown in Figure 6.

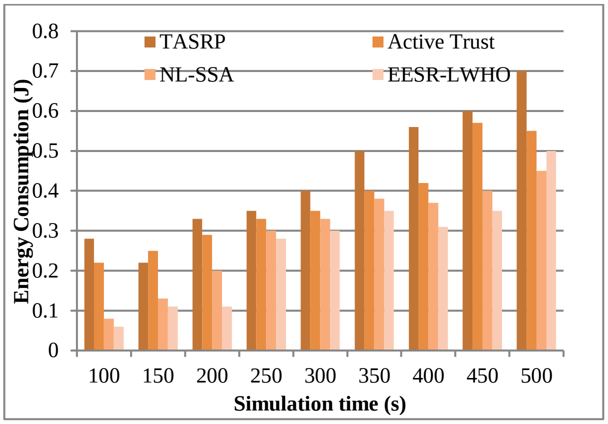

Figure 6: Energy Consumption Analysis

Send/receive cluster head election packets are employed to determine energy consumption by analysing the average power consumption ratio of sensor nodes. 100 sensor nodes are haphazardly placed in a network field the size of a square to conduct a sensitivity evaluation. Moreover, all of the nodes remain stationary throughout the tests, maintaining their placements. The starting energy level of the homogeneous nodes is set to 5 J. Additionally, each node's transmission power is set at 20 m.

V. CONCLUSION

The goal of the research is developing the EESR protocol and node localization using modified wild horse optimization for wireless sensor network-based IoT intrusion protection. The majority of the current solutions ignored intrusions in an unsupervised, infrastructure-less setting by employing a greedy algorithm to design the routing path. When there are many malicious nodes and a high network load, which resulting in a notable quantity of route discoveries and retransmissions. The enhanced method for enhanced localization reconstructs the solution to promptly locate the unidentified sensor node. The randomly deployed electrified node point causes nodes with known location points to be randomly distributed throughout an identical WSN. Furthermore, the ESSR protocol employed a lightweight secret sharing technique among cluster heads and BS to establish a secure network-wide data routing over hostile nodes. This provides data security against malicious threats from nodes to cluster chiefs and beyond to the BS. In summary, when compared to alternative methods, the suggested EESR-LMWHO works well with respect to real position point and optimal security with regard to location. In future study, multi-hop network communication and mobility standards will be taken into consideration while expanding the suggested protocol.

REFERENCES

[1] B.D. Deebak., and F. Al-Turjman, "A hybrid secure routing and monitoring mechanism in IoT-based wireless sensor networks", *Ad Hoc Networks*, Vol. 97, 2020.

[2] S.W. Nourildean., M.D. Hassib., and Y.A. Mohammed, "Internet of things based wireless sensor network: a review", *Indonesian Journal of Electrical Engineering and Computer Science*, Vol. 27, No. 1, Pp. 246-261, 2022.

[3] A. Ghani., K. Mansoor., S. Mehmood., S.A. Chaudhry., A.U. Rahman., and M. Najmus Saqib, "Security and key management in IoT-based wireless sensor networks: An authentication protocol using symmetric key", *International Journal of Communication Systems*, Vol. 32, No. 16, 1-20, 2019.

[4] H. Cheng., C. Rong., and G. Yang, "Design and analysis of a secure routing protocol algorithm for wireless sensor networks", *In IEEE International Conference on Advanced Information Networking and Applications*, Pp. 475-480, 2011.

[5] S. Sharma, *Energy-efficient secure routing in wireless sensor networks* (Doctoral dissertation), 2009.

[6] H.W. Ferng., and D. Rachmarini, "A secure routing protocol for wireless sensor networks with consideration of energy efficiency", *In IEEE network operations and management symposium*, Pp. 105-112, 2012.

[7] M.A. Al-Jarrah., M.A. Yaseen., A. Al-Dweik., O.A. Dobre., and E. Alsusa, "Decision fusion for IoT-based wireless sensor networks", *IEEE internet of things journal*, Vol. 7, No. 2, Pp. 1313-1326, 2019.

[8] J.N. Al-Karaki, "Analysis of routing security-energy trade-offs in wireless sensor networks", *International Journal of Security and Networks*, Vol. 1, No. 3-4, Pp. 147-157, 2006.

[9] J. Sen., and A. Ukil, "A secure routing protocol for wireless sensor networks", *In Computational Science and Its Applications–ICCSA: International Conference, Fukuoka, Japan, Proceedings, Part III*, Vol. 10, Pp. 277-290, 2010. Springer Berlin Heidelberg.

[10] B. Bhushan., and G. Sahoo, "A comprehensive survey of secure and energy efficient routing protocols and data collection approaches in wireless sensor networks", *In IEEE International Conference on Signal Processing and Communication (ICSPC)*, Pp. 294-299, 2017.

[11] H. Shahid., H. Ashraf., H. Javed., M. Humayun., N.Z. Jhanjhi., and M.A. AlZain, (2021). Energy optimised security against wormhole attack in iot-based wireless sensor networks. *Comput. Mater. Contin*, Vol. 68, No. 2, Pp. 1967-1981.

[12] Q. Shi., L. Qin., Y. Ding., B. Xie., J. Zheng., and L. Song, "Information-aware secure routing in wireless sensor networks", *Sensors*, Vol. 20, No. 1, Pp. 1-21, 2019.

[13] T. Khan., and K. Singh, "TASRP: a trust aware secure routing protocol for wireless sensor networks", *International Journal of Innovative Computing and Applications*, Vol. 12, No. 2-3, Pp. 108-122, 2021.

[14] M. Selvi., K. Thangaramya., S. Ganapathy., K. Kulothungan., H. Khannah Nehemiah., and A. Kannan, "An energy aware trust based secure routing algorithm for effective communication in wireless sensor networks", *Wireless Personal Communications*, Vol. 105, Pp. 1475-1490, 2019.

[15] T. Kalidoss., L. Rajasekaran., K. Kanagasabai., G. Sannasi., and A. Kannan, "QoS aware trust-based routing algorithm for wireless sensor networks", *Wireless Personal Communications*, Vol. 110, Pp. 1637-1658, 2020.

[16] Y. Liu., M. Dong., K. Ota., and A. Liu, "Active Trust: Secure and trustable routing in wireless sensor networks", *IEEE Transactions on Information Forensics and Security*, Vol. 11, No. 9, Pp. 2013-2027, 2016.

[17] S. Sivasakthiselvan., and V. Nagarajan, "A new localization technique for node positioning in wireless sensor networks", *Cluster Computing*, Vol. 22, Pp. 4027-4034, 2019.

[18] R.V. Kulkarni., G.K. Venayagamoorthy., and M.X. Cheng, "Bio-inspired node localization in wireless sensor networks", *In IEEE International Conference on Systems, Man and Cybernetics*, Pp. 205-210, 2009.

[19] S. Arora., and S. Singh, "Node localization in wireless sensor networks using butterfly optimization algorithm", *Arabian Journal for Science and Engineering*, Vol. 42, Pp. 3325-3335, 2017.

[20] H.M. Kanoosh., E.H. Houssein., and M.M. Selim, "Salp swarm algorithm for node localization in wireless sensor networks", *Journal of Computer Networks and Communications*, 2019.

[21] J. Cheng., and L. Xia, "An effective cuckoo search algorithm for node localization in wireless sensor network", *Sensors*, Vol. 16, No. 9, Pp. 1-17, 2016.

[22] S. Goyal., and M.S. Patterh, "Modified bat algorithm for localization of wireless sensor network", *Wireless Personal Communications*, Vol. 86, Pp. 657-670, 2016.

[23] I. Naruei., and F. Keynia, "Wild horse optimizer: A new meta-heuristic algorithm for solving engineering optimization problems", *Engineering with computers*, Vol. 38 No. Suppl 4, Pp. 3025-3056, 2022.