Journal of Information Systems Engineering and Management
2025, 10(9s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Machine Learning-Enhanced Hybrid Source Location Privacy Protocol for Improved Security and Network Longevity in IoT Networks

Mrs. Neha Gharat[1, 2*], Dr. Lochan Jolly[3]

[1] Research scholar, Dept. of EXTC, Thakur College of Engg. & Tech., Mumbai University, Maharashtra, India
[2] Assitant Professor, Dept. of EXTC, Vidyavardhini's College of Engg. & Tech., Mumbai University, Maharashtra, India*
[3] Prof., Dept. of EXTC, Thakur College of Engg. & Tech., Mumbai University, Maharashtra, India
**Corresponding email:neha.gharat@vcet.edu.in*

ARTICLE INFO

Received: 29 Oct 2024

Revised: 20 Dec 2024

Accepted: 30 Dec 2024

ABSTRACT

Efficient routing in IoT networks is critical for optimizing data transmission while addressing the inherent challenges of energy consumption, scalability, and resilience. Common routing techniques, such as flooding, tree-based, cluster-based, and geographic routing, each present unique advantages and limitations regarding energy efficiency, network congestion, and fault tolerance. In parallel, source location privacy protection (SLPP) has emerged as a pivotal concern in IoT applications like surveillance and environmental monitoring, where adversaries may exploit transmission patterns to identify sensitive source nodes. Despite advancements, current issues in SLP in IoT networks include balancing privacy with energy efficiency, ensuring scalability in dense network environments, and providing resilience against increasingly sophisticated adversarial models. Traditional SLP techniques, including phantom routing, random walks, and dummy packet generation, often impose trade-offs between privacy, energy consumption, and network longevity, limiting their practical application in large-scale IoT networks. Additionally, many existing protocols struggle with adapting to dynamic network topologies and fail to adequately address the challenges posed by hotspot formation, which can lead to uneven energy depletion and compromised privacy. Many SLP methods disrupt the Quality of Service (QoS) by introducing delays or reducing throughput, which can hinder the primary functions of IoT applications, particularly in time-sensitive scenarios. A new Hybrid Source Location Privacy (SLP) protocol that effectively integrates random walks, rumor routing, and Greedy Random Walks to obscure source node locations while optimizing energy consumption is implemented to overcome these limitations. The protocol employs a multi-layer grid framework, dynamic cluster head rotations, and phantom nodes to balance energy usage and reduce network hotspots. The new Hybrid SLP confuses adversaries by combining fake packet generation with adaptive routing strategies, enhancing privacy without compromising network performance. Simulations demonstrate that the Hybrid SLP protocol significantly outperforms existing techniques, achieving lower energy consumption, extended network lifetime, and robust privacy protections, making it ideal for privacy-sensitive IoT applications. The proposed Hybrid SLP protocol integrates a machine learning-based anomaly detection system to enhance its performance and security, highlighting the novelty of the proposed work. This novel combination of advanced routing strategies and machine learning strengthens network resilience against various threats.

Keywords: Machine Learning, Hybrid Source, IoT

1. INTRODUCTION

The Internet of Things (IoT) is a transformative technology designed to collect, transmit, and utilize data through wireless connections, revolutionizing information access across various fields, including environmental monitoring, healthcare, and smart cities [1]. While IoT deployment is straightforward in areas with stable internet access, challenges arise in remote and resource-constrained environments like dense forests, mountainous regions, and deep-sea research zones. Wireless Sensor Networks (WSNs) bridge this connectivity gap by enabling scalable, flexible deployment for tracking valuable assets, such as real-time military operations and endangered species monitoring. However, direct communication between IoT nodes and the sink node quickly depletes node energy, making energy management critical in IoT applications. Existing routing schemes, including flooding, tree-based, cluster-based, and geographic routing, provide diverse solutions for data transmission but often face trade-offs between energy consumption, fault tolerance, and network congestion, particularly in large-scale networks [2]. Additionally, WSNs are inherently vulnerable to security threats due to the open nature of wireless communication channels, exposing them to risks like eavesdropping and packet dropping.

Traditional Source Location Privacy (SLP) techniques, such as phantom routing, dummy packet generation, and random walks, aim to protect the source node by obscuring data origins [3]. However, these methods frequently introduce high energy overhead, increased latency, and network complexity, limiting their effectiveness in resource-constrained settings. Recent innovations, including tree-based diversionary paths using phantom nodes [10] and multi-sink node routing with dynamic packet destinations [11], offer improved safety periods by misleading attackers. To further address challenges in network longevity and privacy, researchers have explored optimal power management [11], clustering techniques [12], and advanced energy-efficient protocols [13,14]. This paper introduces a novel hybrid approach to SLP, combining advanced privacy protection with extended network lifespan by optimizing energy efficiency and load distribution. The proposed Hybrid SLP method significantly outperforms traditional techniques and is ideal for privacy-critical IoT applications, striking an effective balance between security and energy management in modern networks.

This paper is organized as follows: Section 2 reviews related studies and identifies gaps in existing SLP techniques. Section 3 presents the Hybrid SLP protocol, its key components, and the models used, while Section 4 discusses simulation results, performance metrics, and key findings. Section 5 concludes with a summary and future research directions.

2. LITERATURE SURVEY

Researchers have developed various routing strategies to address Source Location Privacy (SLP) challenges in IoT-enabled Wireless Sensor Networks (WSNs). Baseline flooding enhances privacy by involving all nodes in data transmission, potentially confusing adversaries, but it suffers from high energy consumption and minimal safety margins as packets travel the shortest path to the base station (BS). Attackers can easily trace the source under this approach. To address this, probabilistic flooding selectively involves nodes in the transmission process, reducing energy consumption while maintaining privacy protections. To enhance source location privacy (SLP) in Wireless Sensor Networks (WSNs), researchers have explored various routing strategies, focusing on balancing privacy and resource efficiency. Fake packet-based SLP techniques, extensively examined in studies such as [8, 9, 10, 11-13, 14, 15], introduce dummy packets or fake sources to obscure real transmissions. However, these approaches are often energy-intensive, making them less viable for WSNs with limited resources. Instead, random walk- and phantom routing-based techniques are more suitable for energy-constrained environments [16, 18, 21, 27]. For a comprehensive analysis of SLP preservation strategies, the work by Conti et al. [23] provides valuable insights.

Random walk-based methods improve privacy by creating unpredictable packet paths, complicating adversaries' tracking attempts through random relay node selection. Ozturk et al. [28] introduced the Phantom Flooding Scheme (PFS), combining an initial directed walk with baseline flooding to reach the base station. While PFS enhanced privacy, its uneven node selection weakened protection and failed to optimize network longevity. This highlights the need for improved methods that balance privacy, energy efficiency, and network sustainability. Phantom node-based strategies employ techniques such as phantom single-path routing, locational angle-based phantom routing, phantom walkabouts, two-level phantom with backbone routing, pseudo-normal distribution-based phantom routing, Greedy Random Walk routing, and probabilistic routing [20], [31]. Similarly, angle-based methods encompass protocols like angle-based intermediate node routing, angle-strategic routing, dynamic angle-based routing, angle-proxy routing, constrained random routing, and two-phantom angle-based routing [12], [20].

SLP protocols are categorized into tree-based, intermediate node-based, phantom node-based, and angle-based strategies. Tree-based methods include diversionary and bidirectional tree routing [10], while intermediate node-based approaches involve techniques like randomly selected intermediary node routing and sink toroidal region routing [10], [35]. Hybrid protocols often combine multiple strategies, such as integrating phantom routing with ring and fake packet routing [10], [16]. Proxy nodes also enhance privacy by acting as intermediaries between the source and destination [21]. Multi-sink node protocols [14], [27] and two-phase routing using virtual nodes, escape angles, and random walks further improve security [27]. To address limitations in SLPDR, Ring-Loop Routing (SLPRR) employs phantom nodes, obfuscated transmissions, and random routing for enhanced privacy. Comparative evaluations indicate that no single protocol effectively balances privacy, energy efficiency, and network longevity, emphasizing the need for hybrid SLP solutions.

3. PROPOSED METHODOLOGY

The new Hybrid SLP protocol introduces an innovative technique that enhances both the safety period and network longevity in IoT-enabled WSNs. It improves upon traditional random walk-based SLP methods by offering superior privacy protection and extended network lifespan, even in large-scale or challenging environments. By balancing privacy and energy efficiency, it is ideal for resource-constrained IoT networks. The Hybrid SLP method is designed to safeguard the source node's location is as shown in Figure 1. The new Hybrid SLP routing protocol effectively combines random walks, (RW) rumor routing, and Greedy Random Walks to achieve strong source location privacy, efficient energy usage, and scalability in IoT networks. Its ability to obscure the source node's location while optimizing real packet delivery makes it a robust solution for privacy-sensitive and resource-constrained IoT applications. IoT nodes are deployed, and the network is divided into layers and grids for efficient organization.

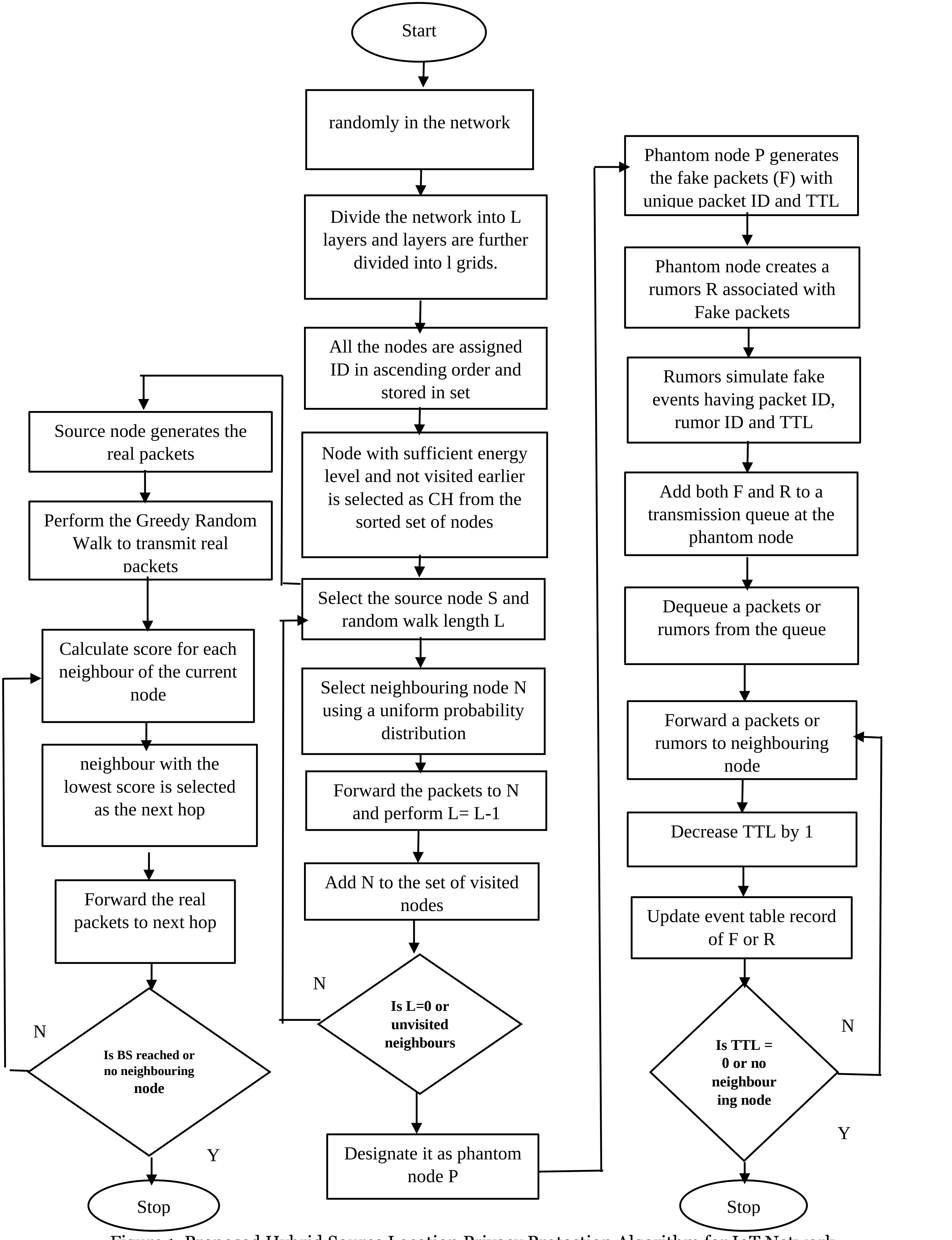

Start
randomly in the network
Divide the network into L layers and layers are further divided into l grids.
All the nodes are assigned ID in ascending order and stored in set
Node with sufficient energy level and not visited earlier is selected as CH from the sorted set of nodes
Select the source node S and random walk length L
Select neighbouring node N using a uniform probability distribution
Forward the packets to N and perform L= L-1
Add N to the set of visited nodes
Is L=0 or unvisited neighbours
N
Designate it as phantom node P
Source node generates the real packets
Perform the Greedy Random Walk to transmit real packets
Calculate score for each neighbour of the current node
neighbour with the lowest score is selected as the next hop
Forward the real packets to next hop
Is BS reached or no neighbouring node
N
Y
Stop
Phantom node P generates the fake packets (F) with unique packet ID and TTL
Phantom node creates a rumors R associated with Fake packets
Rumors simulate fake events having packet ID, rumor ID and TTL
Add both F and R to a transmission queue at the phantom node
Dequeue a packets or rumors from the queue
Forward a packets or rumors to neighbouring node
Decrease TTL by 1
Update event table record of F or R
Is TTL = 0 or no neighbour ing node
N
Y
Stop

Figure 1. Proposed Hybrid Source Location Privacy Protection Algorithm for IoT Network

A random walk is initiated from the source node to identify a phantom node. This technique selects neighbors probabilistically over a predefined number of hops, creating a randomized path that obscures the true location of the source. The final node reached after the random walk is designated as the phantom node, which serves as an intermediate point for routing packets, adding an extra layer of privacy. At the phantom node, fake packets and rumors are generated. Fake packets simulate decoy traffic, while rumors mimic event-driven communication, such as alerts or network status. These are propagated using rumor routing, where packets are forwarded probabilistically to a subset of neighboring nodes. This generates trails of decoy traffic, complicating efforts for adversaries to pinpoint the actual source location. Genuine packets are sent from the origin node to the base station through a Greedy Random Walk strategy.

Random walks and rumor routing create multiple decoy paths, protecting the source node's location from adversaries. Greedy Random Walks balance energy consumption across the network, preventing nodes from being overburdened and extending the network lifetime. The layered and grid-based structure adapts well to large IoT networks, ensuring efficient communication in dense deployments. Rumor routing minimizes unnecessary network-wide communication by limiting packet propagation to a subset of nodes.

Network Model:

The network consists of 200 randomly or grid-deployed nodes using multi-hop routing to forward data to a single sink node. Rumor Routing and Greedy Random Walk are used to create unpredictable, energy-efficient paths while avoiding hotspots. Fake packets are generated and routed along random or greedy paths to obscure the true source node’s location and mislead attackers. Nodes with limited initial energy E0 balance load distribution, ensuring extended network longevity and enhanced privacy protection.

i. Attacker Model:

A cautious adversary with ample resources, powerful transceivers, and sufficient storage can detect packet signals and monitor traffic patterns. They can infer the transmitter's location by analyzing traffic converging at the central hub and backtracking to the source. Additionally, they can analyze captured packets for routing patterns and compromise nodes to extract or manipulate information. However, the adversary's mobility is limited, restricting their ability to monitor the entire network simultaneously.

ii. Cluster Head (CH) Rotation:

IoT nodes are arbitrarily positioned inside the network area with coordinates (xi, yi) for each node i. This approach utilizes a multi-layer framework, where each layer is subdivided into multiple grids. The network region is organized into L layers, each further divided into $l$grids. Each layer and grid is assigned a unique identifier to facilitate organization within the network. It is assumed that all grids and layers have equal areas. The neighboring nodes of a given node are denoted as: N(i) = {j:(i, j)∈E}, representing the set of nodes directly connected to node i. The base station (BS) is located at the apex of the network, while a phantom node is chosen through a random walk process. Each grid contains a node assigned as the cluster head (CH). The Cluster Head (CH) role assignment is a dynamic process that plays a crucial role in balancing the communication load among IoT nodes in a network. Within each grid, a node is designated as the CH to manage local communication, coordinate data aggregation, and forward packets to the base station (BS) or neighboring grids. The nodes take turns handling this responsibility in a cyclic order after each packet transmission. The rotation ensures that no single node is overburdened, distributing the energy consumption evenly along the network.

CH role is updated using the equation 1:

$$\text{Next CH} = (\text{Current CH Index} + 1) \bmod \mid \mathrm{N_{grid}} \mid \qquad (1)$$

Where,

$|N\text{grid}|$ is the number of nodes within the grid.

All nodes start with equal energy:

$$\mathrm{E_{init(i)}} = \mathrm{E_{total}}/|\mathrm{N_{grid}}\ |,\ \forall \mathrm{i} \in \mathrm{N_{grid}} \qquad (2)$$

The energy consumed by the CH in Intra-grid communication is given by

$$\mathrm{E}_{comm} = \mathrm{P_{tx}} \cdot \mathrm{d} + \mathrm{P}_{rx} \cdot \mathrm{n} \qquad (3)$$

where:

P_{tx}: Transmission power,

P_{rx}: Reception power,

d: Distance to neighboring nodes,

n: Number of packets transmitted or received

The CH forwards aggregated data to another CH or the BS. The energy consumed by the CH in Inter grid communication is given by equation 4.

$$E_{forward} = Ptx \cdot D_{grid-to-BS} \qquad (4)$$

Where $D_{grid-to-BS}$ is the distance from the CH to the BS

The total energy consumed by a CH per transmission round is given by equation 5.

$$E_{CH} = \mathrm{E}_{comm} + E_{forward} \qquad (5)$$

After one round the residual energy of a node $i \in N$grid is given by equation 6.

$$E_{res(i)} = \begin{cases} E_{init}\ (i) - E_{CH},\ if\ \ i\,was\,CH \\ \qquad E_{init}\ (i),\ \ otherwise \end{cases} \qquad (6)$$

Residual energy across all nodes in the grid after R rounds:

$$E_{res}\ (i,R) \approx\ E_{init}\ (i) - \frac{R}{|N_{grid}|} * E_{CH} \qquad (7)$$

Nodes with residual energy $E_{res}(i,\ R)$ below $E_{threshold}$ are excluded from the CH rotation. The CH role rotation is adjusted dynamically to balance the load among remaining nodes:

R_{CH} (j) is updated for all j ∈ N_{grid} \ {i: E_{res} (i, R) < $E_{threshold}$}

However, nodes with residual energy below a predefined threshold, Ethreshold are excluded from CH selection to preserve their energy for other network functions. This approach ensures that nodes with sufficient energy levels are prioritized, preventing premature node depletion

iii. Phantom Routing:

The primary objective is to protect the confidentiality of the transmitting node by utilizing random routing paths for packet delivery and incorporating decoy packet transmission from phantom nodes (PNs). The base station (BS) is located at the apex of the network, while a phantom node is chosen through a random walk process. Each grid contains a node assigned as the cluster head (CH). The process of selecting a phantom node in Phantom Routing involves creating a diversion from the actual source node's location to obscure its identity and protect its privacy. A random walk is initiated from the source node to select a phantom node. The phantom node P is selected using a random walk from the source node S for T phantom steps. At each step t, the probability of moving from the active node i to a neighboring node j is given by equation (8).

$$Pt(i \rightarrow j) = \begin{cases} \frac{1}{|N(i)|},\ if\ j \in N(i) \\ \quad 0,\ if\ j\ \notin N(i) \end{cases} \qquad (8)$$

where

N(i) is the set of neighbors of node u,

Rumor Routing:

Rumor Routing and Phantom Routing together enhance Source Location Privacy (SLP) by leveraging the creation of trails (Rumor Routing) and the randomized selection of phantom nodes (Phantom Routing). Once the phantom node P is selected, the protocol leverages Rumor Routing to forward packets efficiently to the BS. Rumors carry the contextual information necessary to propagate fake packets through the network. For example, a rumor may indicate a simulated event ("Anomaly detected at Node X") or create a fake route to guide fake packets along unpredictable paths. In Rumor Routing, mobile agents are deployed to create trails between CHs and the BS. These agents record state information at each visited node, forming trails. These agents traverse the network using random walks. Mobile agents also record state information T (i, j) at each visited node i for edge (i, j)

$$\mathrm{T\ (i, j)\ =\ True}, \forall\ \mathrm{(i, j)}\ \in\ \mathrm{trail} \qquad (9)$$

Trails are virtual paths in the network that connect CHs to the BS or other CHs. These trails form paths from CHs to the BS, ensuring efficient data delivery. Trails can expire or be updated dynamically based on network conditions, such as node failures or energy depletion. A packet originating at a phantom node (P), follows the shortest trail R (P, BS) based on state information as follows

$$R(P, BS) = \min \textstyle\sum_{(i,j) \in \mathrm{T}} d(i, j) \qquad (10)$$

where

d(i, j) is the distance between nodes i and j.

When a node i receives a packet, it uses its routing table T(i, j) to forward the packet along a valid trail to the BS. The next hop i is selected based on the trail's state information given by expression as

$$j\,j = \arg \underset{i\ \in\ \mathrm{N\ (i)}}{min}\ T(i, j) * d(i, j) \qquad (11)$$

iv. Greedy Random Walk:

Greedy Random Walk enhances the hybrid protocol by optimizing forwarding decisions along the trails established by Rumor Routing. Unlike static routing, where paths are pre-defined, Greedy Random Walk dynamically evaluates neighboring nodes at each step based on their residual energy and distance to the BS. The goal of Greedy Random Walk is to select the next hop along the trail in a way that optimizes energy consumption and reduces latency. While the Greedy Random Walk dynamically optimizes forwarding decisions, it still adheres to the trails established by Rumor Routing. This ensures efficient use of established paths while allowing flexibility in node selection. Nodes with sufficient energy along the trail are prioritized, preventing the depletion of specific nodes and prolonging the network's lifetime. Nodes with sufficient energy along the trail are prioritized, preventing the depletion of specific nodes and prolonging the network's lifetime. Nodes with higher E (j) are preferred to avoid depleting energy of critical nodes. The Euclidean distance between a node j and the BS. Nodes closer to the BS are preferred to minimize latency and transmission energy. For the current node i, the neighbors N(i) are evaluated based on the scoring function. Each neighbor i receives a score based on its energy and proximity to the BS. The node with the highest score is selected as next hop i*. The packet is forwarded to i*, and the process repeats at i* until BS is reached. If two nodes have similar scores, tie-breaking rules (e.g., random selection) prevent overloading a single node. The score combines distance to the base station and, optionally, other factors like randomization or energy levels to ensure privacy and efficiency. At each step, the next hop i is selected based on a scoring function:

$$Score(j) = \alpha * \frac{E(j)}{max_{\omega \in N(i)} E(\omega)} - \beta * \frac{d(i,BS)}{max_{\omega \in N(i)} E(\omega, BS)} \qquad (12)$$

where:

i: Current node.

N(i): Neighbouring nodes of i.

α, β: Weighting factors for energy and distance.

maxω∈N(i) E (ω) : maxω∈N(i)d (ω, BS): Normalize the values to ensure fair comparison.

The density and distribution of nodes affect the path length. A dense network reduces the average distance between nodes, minimizing path length. The total path length Lpath from the source node S to the base station BS is the sum of the contributions from Phantom Routing and Rumor Routing, optimized by Greedy Random Walk. The total path length is given by

$$L_{path} = L_{phantom} + L_{greedy}$$

Greedy Random Walk reduces the number of hops in Lrumor by prioritizing closer nodes.

$$L_{greedy} \leq L_{rumor}$$

By dynamically balancing energy consumption and minimizing latency, it ensures efficient routing while maintaining robust Source Location Privacy. This combination allows the protocol to adapt to varying network conditions, making it ideal for scalable and privacy-critical sensor networks.

v. Application of Machine Learning alogritm to Detect Sybil and Sinkhole Attacks:

The most detrimental form of routing attack in IoT networks are the Sybil attack and Sinkhole Attacks. In Sinkhole Attack, compromised node lures nearby traffic by falsely advertising itself as the optimal path to the base station, causing data loss, packet dropping, and network performance degradation. In Sybil Attack, a malicious node in the network creates multiple fake identities to deceive legitimate nodes. Isolation Forest can effectively detect both Sybil and Sinkhole attacks in IoT networks. Its energy efficiency, unsupervised nature, and scalability make it well-suited for anomaly detection in resource-constrained, large-scale IoT deployments. The poroposed Hybrid SLP is tested agaist these two attacks by using Isolation Forest algorithm. During the Pre-routing Phase, the Isolation Forest continuously monitors network traffic and detects anomalies during the random walks and Greedy Random Walk processes. As part of Dynamic Adaptation, the protocol adjusts routing strategies in real time based on the detection results, ensuring compromised nodes are either rerouted or dropped. Simultaneously, Energy Optimization is achieved by isolating nodes with high anomaly scores, preventing unnecessary resource depletion and enhancing the network's longevity. The isolation forest anomaly score is calculated using the path length of points through the trees.

$$\text{Isolation score: } s(X_i) = 2^{-\frac{E(h(X_i))}{c(n)}} \qquad (13)$$

Where:

$h(X_i)$is the path length of data point X_i

E(h(Xi) is the average path length across all trees.

c(n) is the normalization factor for the expected path length.

The isolation forest is trained on normal traffic data (without attacks). The model is fitted by using the following loss function:

$$Loss = \sum_{i=1}^{N} \max(0, s(X_i) - Threshold \qquad (14)$$

If the anomaly score s(Xi)s(X_i)s(Xi) exceeds the threshold, the point is flagged as suspicious

4. RESULT AND DISCUSSION

The simulation of the SLP method was conducted using Python on 64 bit Microsoft Windows System, 8 GB RAM, Intel core CPU running at 1.8GHz. The parameters are outlined in Table 1. The simulations are performed 15 times, each with a distinct random deployment of nodes to ensure robustness and reliability of results. The results are validated through comparisons with state-of-the-art approaches, confirming the robustness of the proposed framework.

Table 1. Simulation parameters

Parameters	Values
Number of IoT nodes (N)	200
Network region (M×M)	200 × 200 m^2
E_0 (Initial energy of deployed node)	0.5 J
E_{elec} for all the nodes	50 nJ/bit
ϵfs denotes the energy consumption in the free space model.	10 pJ/bit/m2
ϵ_{mp} represents the energy consumption in the multi-path model	0.0013 pJ/bit/m4
L (packet size in bits)	1000 bits
Sensor node transmission range	40 m

The simulation in Figures 2 illustrates the Hybrid Source Location Privacy (SLP) Protocol in an IoT network, showing the transmission of real and fake packets through a 200×200-meter 2D space. The real packet (red) travels from the source node (yellow) through phantom nodes (red), cluster heads (green), and reaches the destination node (purple). Fake packets (orange) follow randomized paths through various phantom nodes to confuse adversaries and enhance privacy. The simulation highlights the Hybrid SLP protocol's effectiveness in bal ancing privacy, energy efficiency, and path diversity within the network.

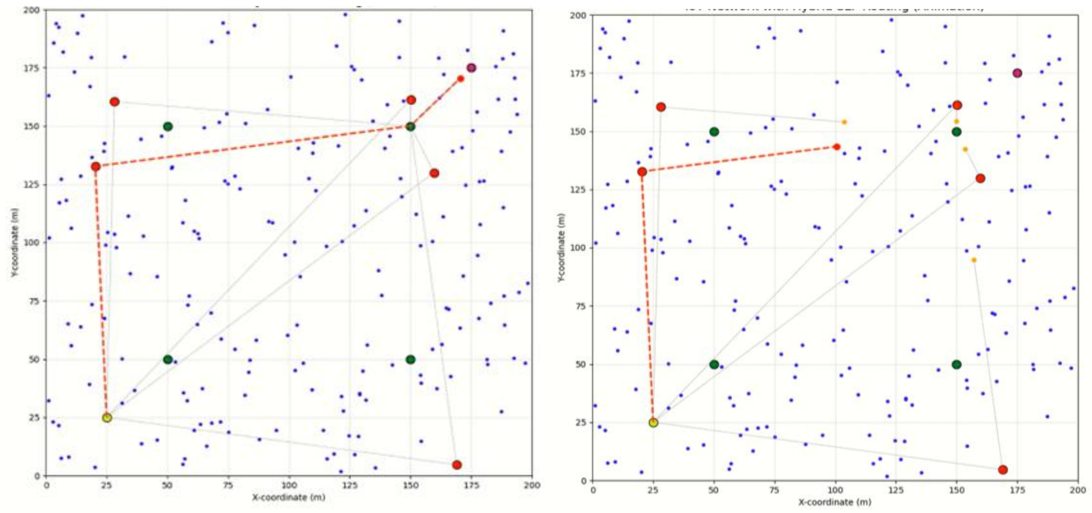


Figure 2: Simulation of packet transmission in an IoT network using the Hybrid Source Location Privacy (SLP) Protocol

The effectiveness of the proposed technique is assessed using the following performance metrics.

1. Energy consumption: The energy consumed during the transmission of the l-bit packet over distance d is represented by E_{Tx}(l,d).

$$E_{Tx} = 1 * \left(E_{elec} + E_{freespace} * d^2\right) if\ d < d_0 \qquad (12)$$

$$E_{Tx} = 1 * \left(E_{elec} + E_{multipath} * d^4\right) if\ d \geq d_0 \qquad (13)$$

The energy consumed during the reservation of the l-bit packet is given by equation 14.

$$E_{rx} = 1 * E_{elec} \qquad (14)$$

In the above energy equation, E_{elec} represents the energy expended by the transmitter or receiver circuitry, while the energy required by the transmission amplifier for free space and multipath propagation is denoted by E_{fs} and E_{mp}. The threshold value d_0 is $\sqrt{E_{freespace/E_{multipath}}}$

2. Network Lifetime: This refers to the period from when the network begins operating until the first node's energy is completely depleted. This is determined by calculating the total number of messages successfully delivered to the base station (BS) prior to the depletion of energy in the first node.

3. Transmission Delay: The Source Location Privacy (SLP) protocol ensures route privacy by assigning each packet a unique path originating at the source node and terminating at the base station (BS). To evaluate delay performance, the hop count for each packet traveling between the source node and the BS is measured. Greater variability in hop counts across different routes enhances the overall privacy of the transmission.

A. Energy Consumption:

The proposed Hybrid SLP protocol improves source node privacy by using dummy packet transmissions through phantom nodes, which disrupt attacker backtracking attempts and are primarily generated by IoT nodes in the hotspot area. This method involves more dummy packets than existing schemes but is considered acceptable due to the enhanced privacy it provides. Energy consumption data from Table 2 shows that the Baseline Protocol has the highest energy use across all distances, near 400 units, due to its lack of advanced privacy mechanisms. The Phantom and Probabilistic protocols consume slightly less energy, while the SLPDR, SLPRR, and SLP-RRFPR protocols show significantly lower and more consistent energy consumption, below 50 units. The Proposed SLP protocol is the most energy-efficient, with consistently lower energy consumption compared to all other protocols, indicating superior privacy and optimized resource utilization.

Table 2. Comparison of Energy consumption versus distance from the SN to BS for various SLPs

Distance between SN and BS (Meters)	Baseline [28]	Phantom [29]	Probabilistic [30]	SLPDR [2]	SLPRR [2]	SLP-RRFPR [3]	Proposed Hybrid SLP
20	390	380	380	10	10	20	4
40	390	380	380	12	10	20	8
60	390	380	350	12	10	20	8
80	390	380	380	12	12	22	8
100	390	380	380	14	15	25	12
120	390	380	380	25	20	25	18
140	390	380	380	32	27	25	18
160	390	380	380	40	34	25	18

Figure 3: Energy Consumption (μJ) Vs SN to BS Distance

Table 3 compare the energy consumption of various SLP protocols based on communication radius, showing that the Baseline protocol consumes the most energy, around 400 units across all radii. Phantom and Probabilistic protocols use slightly less energy than the Baseline, with Probabilistic routing showing a minor dip at smaller radii. The SLPDR, SLPRR, and SLP-RRFPR protocols exhibit significantly lower energy consumption, consistently staying below 50 units, reflecting more energy-efficient designs. The Proposed Hybrid SLP protocol outperforms all others, maintaining the lowest energy consumption across all communication radii by effectively combining Rumor Routing, random walk, and Greedy Random Walk strategies.

Table 3 Comparison of Energy consumption versus communication radius for various SLPs

Communication radius (Meters)	Baseline [28]	Phantom [29]	Probabilistic [30]	SLPDR [2]	SLPRR [2]	SLP-RRFPR [3]	Proposed Hybrid SLP
20	400	380	390	10	10	25	12
25	400	350	390	12	10	20	14
30	400	390	390	15	12	19	16
35	400	390	390	25	19	19	16
40	400	390	390	40	25	19	16

B. **Network lifetime:**

Table 4 compare the network lifetime of various SLP protocols across different communication radii. The Baseline protocol shows a steady lifetime of about 1000 rounds, while Phantom and SLPDR experience declines at larger radii due to inefficient energy management. SLPRR and SLP-RRFPR maintain stable lifetimes with gradual improvements as the communication radius expands. The Proposed SLP significantly outperforms all other protocols, achieving up to 2500 rounds at a 40-meter radius by effectively balancing energy consumption using Rumor Routing, random walk, and Greedy Random Walk techniques.

Table 4 Comparison of Network lifetime versus communication radius for various SLPs

Communication radius (Meters)	Baseline [28]	Phantom [29]	Probabilistic [30]	SLPDR [2]	SLPRR [2]	SLP-RRFPR [3]	Proposed Hybrid SLP
20	1150	1100	1400	200	100	1700	1250
25	1150	1120	1420	600	50	1500	1550
30	1150	1120	1420	450	200	1450	1666
35	1150	1120	1420	450	400	1400	2000
40	1150	1120	1410	450	480	1400	2500

Table 5 compare the network lifetime of various SLP protocols as a function of the distance between the Source Node (SN) and Base Station (BS). The Baseline protocol shows a steady 1000-round lifetime, while Phantom, Probabilistic, and SLP-RRFPR stabilize around 1500 rounds, with Phantom slightly outperforming the others. SLPDR and SLPRR have significantly lower lifetimes, remaining below 500 rounds, particularly at shorter distances. The Proposed SLP achieves the longest lifetime, peaking at 3500 rounds at shorter distances and gradually decreasing to 2000 rounds, consistently outperforming all other protocols due to superior energy management.

Table 5 Comparison of Network lifetime versus distance from the SN to BS for various SLPs

Distance between SN and BS (Meters)	Baseline [28]	Phantom [29]	Probabilistic [30]	SLPDR [2]	SLPRR [2]	SLP-RRFPR [3]	Proposed Hybrid SLP
20	1150	1100	1400	200	100	1600	3550
40	1150	1120	1420	600	50	1580	3333
60	1150	1120	1420	450	200	1400	2650
80	1150	1120	1420	450	400	1500	2335
100	1150	1120	1410	450	480	1600	2200
120	1150	1120	1420	420	425	1480	2200
140	1150	1120	1420	400	425	1600	2200
160	1150	1120	1420	450	450	1200	2200

The validation of the proposed Hybrid SLP against existing protocols demonstrates superior performance in terms of energy efficiency and network lifetime

C. Transmission delay and safety period:

In Figure 3 comparison of the average packet delay (in number of hops) for multiple Source Location Privacy (SLP) protocols across varying distances between the Source Node (SN) and the Base Station (BS) is given. The analysis encompasses Baseline [28], Phantom [29], Probabilistic [28],, SLPDR, SLPRR [2], SLP-RRFPR [3],, and the Proposed Hybrid SLP Protocol. The Proposed Hybrid SLP protocol achieves the lowest delay at shorter (20-40m) and larger (120-140m) distances, demonstrating superior adaptability and routing efficiency. It dynamically adjusts to network conditions, minimizing hops and avoiding routing loops even at greater distances. The protocol effectively balances source location privacy and routing efficiency, ensuring minimal delays without compromising security. Its scalability and advanced pathfinding enable efficient long-path management, outperforming traditional protocols.

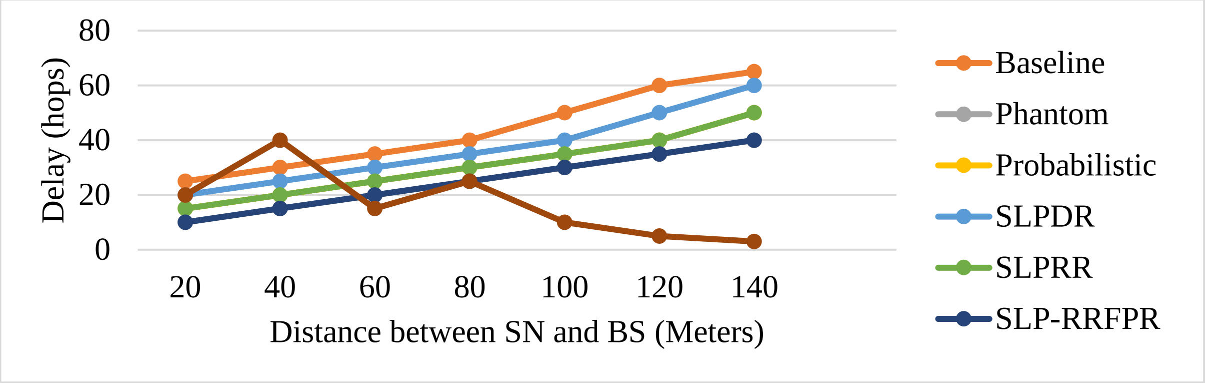


Figure 3: Average packet delay versus distance from the SN to BS

The graph in Figure 4 shows the average delay (in hops) for various protocols across different communication radii. The Baseline protocol consistently has the highest delays, while the Phantom and Probabilistic protocols show moderate improvements but still experience higher delays at larger radii. SLPDR, SLPRR, and SLP-RRFPR demonstrate better efficiency, with SLPRR and SLP-RRFPR maintaining relatively low delays. The Proposed SLP protocol outperforms all others, achieving the lowest delays, particularly excelling in the 20 to 40-meter radius range, highlighting its advanced optimization and efficient routing strategies.

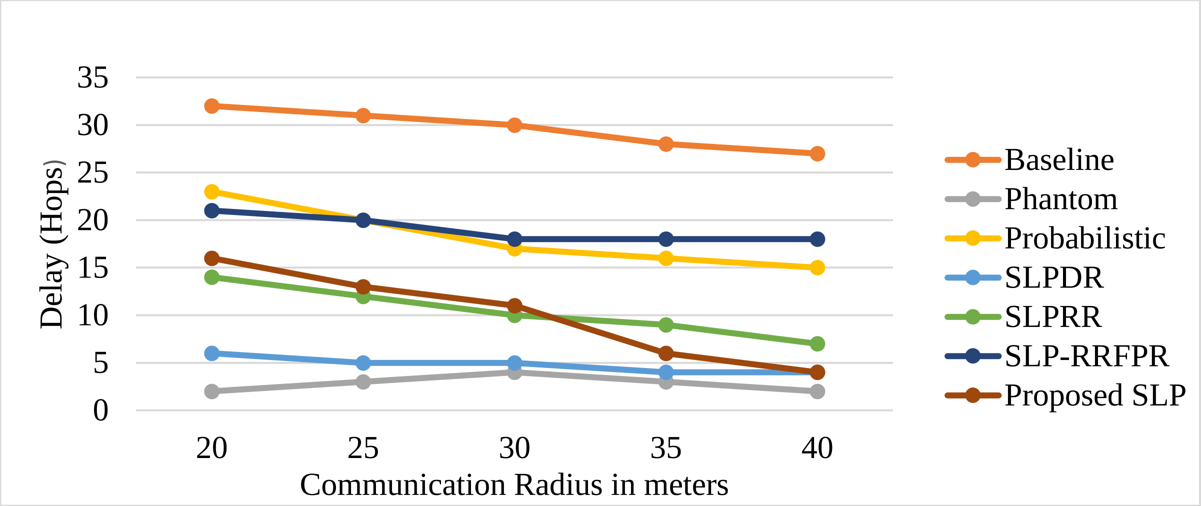


Figure 4 Average packet delay versus varied communication radius

Table 6 highlights the average delay introduced by various Source Location Privacy (SLP) techniques, indicating the network's safety time and privacy protection. The Proposed SLP achieves the highest delays, such as 130 hops at 20 meters and 85 hops at 40 meters, offering superior privacy and extended safety time. SLP-RRFPR provides moderate delays, balancing reasonable privacy with faster transmission, while SLPDR, SLPRR, and Probabilistic protocols prioritize speed with delays ranging between 6 and 22 hops. Baseline and Phantom methods have minimal delays, prioritizing fast delivery but offering limited resistance against adversarial tracing.

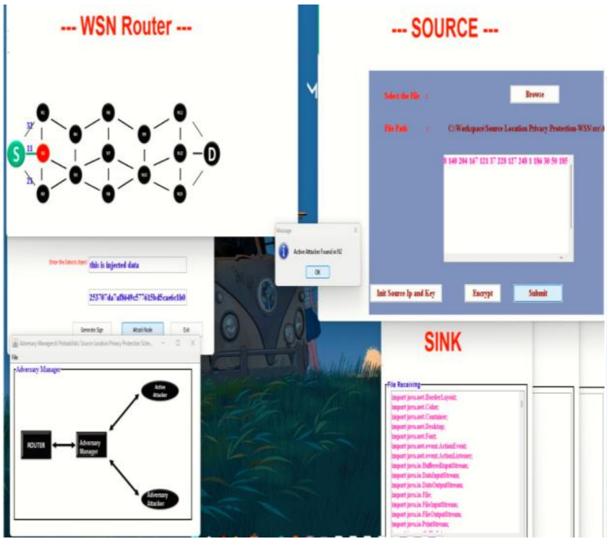
Table 6 Comparison of Average hops by adversary vs Communication Radius for various SLPs

Communication Radius in meter	Baseline [28]	Phantom [29]	Probabilistic [30]	SLPDR [2]	SLPRR [2]	SLP-RRFPR [3]	Proposed Hybrid SLP
20	32	2	22	6	14	60	130
25	31	3	19	5	12	55	105
30	30	4	17	5	10	48	82
35	28	2	16	4	9	46	84
40	26	1	15	4	8	45	85

.

D. Integration of Machine learning with Hybrid SLP:

The use of Isolation Forest within the Hybrid SLP protocol demonstrates robust and scalable detection of Sybil and Sinkhole attacks. This study used IoT-23 datasets, with 25 critical features. After preprocessing and cleaning 48005 samples are used. In the IoT network data is transmitted using hybrid encryption combinig AES and ECC algorithms. Figure 5 and 6 shows attack detection and rerouting using isolation Forest algorithm. When an attacker is detected at node 2, the transmission halts, and the adversary manager evaluates the attack type. Once the attack is indefied the system initiates a rerouting process, directing the packet transmission from the source to the destination through an alternative path.



C:\Workspace\Source Location Privacy Protection-WSN\src\

0 140 204 167 121 37 228 127 248 1 186 30 50 185

Active Attacker Found in N2

Enter the Data to Inject: this is injected data

25370'da"af8649c5"7615bd5ca6c1b0

Adversary Manager & Probabilistic Source Location Privacy Protection Sche...

```
import java.awt.BorderLayout;
import java.awt.Color;
import java.awt.Container;
import java.awt.Desktop;
import java.awt.Font;
import java.awt.event.ActionEvent;
import java.awt.event.ActionListener;
import java.io.BufferedInputStream;
import java.io.DataInputStream;
import java.io.DataOutputStream;
import java.io.File;
import java.io.FileInputStream;
import java.io.FileOutputStream;
import java.io.PrintStream;
```

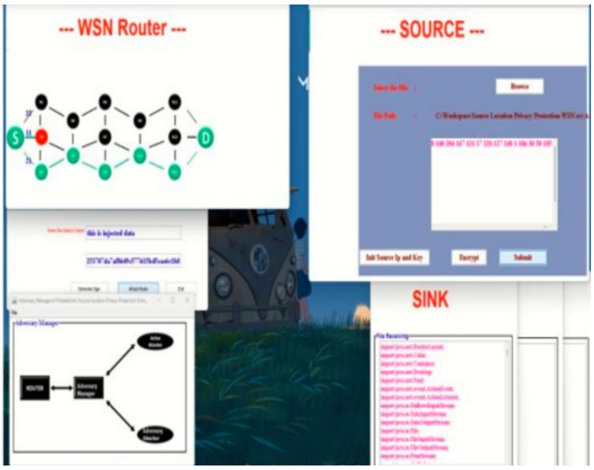


Figure 5. Attack detected at node 2 Figure 6. System initiates a rerouting process

CONCLUSION:

The Proposed Source Location Privacy (SLP) protocol achieves exceptional energy efficiency by utilizing Rumor Routing, which reduces network-wide flooding by involving only a subset of nodes in the routing process. This selective approach creates trails and employs agents to significantly lower communication overhead compared to protocols that depend on extensive network-wide communication. The inclusion of Greedy Random Walk in the Proposed SLP further enhances efficiency by randomizing path selection, preventing certain nodes from becoming overburdened. This helps distribute energy consumption evenly across the network and avoids rapid energy depletion of specific nodes. By integrating Rumor Routing, Random Walk, and Greedy Random Walk, the Proposed SLP achieves minimal energy consumption while ensuring robust source location privacy and a balanced network performance. Across all communication radii, the energy consumption in the Proposed Hybrid SLP method shows consistent improvements of approximately 96% to 97% compared to the Baseline, phantom and probabilistic approach and approximately 20% improvement compared to SLPDR, SLPRR and SLP-RRFPR for larger communication radii.

Although protocols like SLPRR and SLP-RRFPR exhibit reasonable efficiency, they do not match the superior performance of the Proposed SLP. Latency in the hybrid protocol stems from the combined effects of Phantom Routing, Rumor Routing, and Greedy Random Walk. While Phantom Routing's added randomness introduces some delay, the efficiency of Rumor Routing and the optimization from Greedy Random Walk offset this impact. The Proposed SLP achieves a balanced trade-off, delivering strong privacy with manageable delays, making it ideal for scalable, privacy-sensitive applications. The validation of these findings suggests that the approach can be effectively applied in real-world IoT deployment.

The integration of Hybrid SLP with machine learning significantly enhances the protocol's ability to detect and mitigate security threats in dynamic IoT networks. By leveraging machine learning algorithms like Isolation Forest, the system effectively identifies anomalies, such as Sybil and Sinkhole attacks, with high accuracy and minimal energy consumption. This combination ensures adaptive, real-time responses to threats while maintaining network efficiency and extended lifespan

REFRENCES

[1] Shah, S. H., & Yaqoob, " A survey: Internet of things (IoT) technologies, applications, and challenges", *Proceedings of the 2016 IEEE Smart Energy Grid Engineering (SEGE)*, Oshawa, ON, Canada, 21–24 August 2016, 381–385.

[2] Han, G., Zhou, L., Wang, H., Zhang, W., & Chan, S. "A source location protection protocol based on dynamic routing in WSNs for the Social Internet of Things", Future Generation Computer Systems, 82, pp 689–697, 2018. https://doi.org/10.1016/j.future.2017.08.044]

[3] Shukla, A., Tripathi, S., & Singh, K. "SLP-RRFPR: A source location privacy protection scheme based on random ring and limited hop fake packet routing for wireless sensor networks.", *Multimedia Tools and Applications*, 2022. https://doi.org/10.1007/s11042-022-12157-y

[4] Sharma, B., Khosla, A., & Jha, V. "Source location privacy preservation in IoT-enabled event-driven wireless sensor networks", *Wireless Personal Communications, 121*(2), pp.1523–1542, 2021. https://doi.org/10.1007/s11277-021-08978-8

[5] Jimoh, J. B. et.al., "Privacy and security concerns in IoT-based healthcare systems", In *The Fusion of Internet of Things, Artificial Intelligence, and Cloud Computing in Health Care* (pp. 105–134), 2021. Springer: Berlin/Heidelberg, Germany.

[6] Shukla, A., & Tripathi, S. "An effective relay node selection technique for energy-efficient WSN-assisted IoT", *Wireless Personal Communications*, 112(4), pp. 2611–2641, 2020. https://doi.org/10.1007/s11277-020-07167-8

[7] Pai, S., Meingast, et.al., "Transactional confidentiality in sensor networks", *IEEE Security & Privacy, 6*(2), 28–35. [CrossRef]

[8] Wang N., Fu J., Li, B. K., & Bhargava, B. K. "Source-location privacy protection based on anonymity cloud in wireless sensor networks", *IEEE Transactions on Information Forensics and Security, 15*, pp. 100–114, 2019. [CrossRef]

[9] Wan, H., Wu, L., Zhao, Q., Wei, Y., & Jiang, H "Energy balanced source location privacy scheme using multi-branch path in WSNs for IoT", *Wireless Communications and Mobile Computing, 2021*, 6654427. [CrossRef]

[10] Long, J., Dong, M., Ota, K., & Liu, A "Achieving source location privacy and network lifetime maximization through tree-based diversionary routing in wireless sensor networks", *IEEE Access, 2*, 633–651, 2014. [CrossRef]

[11] Chen, H., & Lou, W. "On protecting end-to-end location privacy against local eavesdroppers in wireless sensor networks" *Pervasive and Mobile Computing, Vol.16*, pp. 36–50, 2014. [CrossRef]

[12] Wang, Het.al. "A probabilistic source location privacy protection scheme in wireless sensor networks", *IEEE Transactions on Vehicular Technology, Vol. 68*(7), pp. 5917–5927., 2019 [CrossRef]

[13] Han, G., Wang, H., Miao, X., Liu, L., Jiang, J., & Peng, Y "A dynamic multipath scheme for protecting source-location privacy using multiple sinks in WSNs intended for IIoT", *IEEE Transactions on Industrial Informatics, 16*(8), 5527–5538. [CrossRef]

[14] Han, G., Miao, X., Wang, H., Guizani, M., & Zhang, W "A cloud-based scheme for protecting source location privacy in wireless sensor networks using multi-sinks", *IEEE Transactions on Vehicular Technology, Vol.68*(3), pp. 2739–2750,2019. [CrossRef]

[15] Han,G. Liu, Y. Wang, H. Zhang, Y., "A collision-free-transmission-based source location privacy protection scheme in UASNs under time slot allocation", *IEEE Internet Things*, 2023, vol. 10, pp.1546–1557. [CrossRef]

[16] Mutalemwa a, L. C., & Shin, S., "Achieving source location privacy protection in monitoring wireless sensor networks through proxy node routing", *Sensors, 19*(4), 1037, 2019. [CrossRef]

[17] Zhou, Z., Wang, Y., Li, P., Chang, X., & Luo, J. (2021). Node location privacy protection in unattended wireless sensor networks. *Mathematical Problems in Engineering, 2021*, 5539382. [CrossRef]

[18] Christopher, V., & Jasper, J. "Dynamic routing protocol with mobile sink for location privacy and congestion avoidance in wireless sensor networks", *Journal of Systems Architecture, 112*, 101840, 2021. [CrossRef]

[19] Conti, M., Willemsen, J., & Crispo, B. "Providing source location privacy in wireless sensor networks: A survey", *IEEE Communications Surveys & Tutorials, 15*(3), 1238–1280,2013. [CrossRef]

[20] Jiang, J., Han, G., Wang, H., & Guizani, M. A survey on location privacy protection in wireless sensor networks. *Journal of Network and Computer Applications, 125*, pp. 93–114, 2019. [CrossRef]

[21] He, Y. Han, G. Wang, H. Ansere, J.A. Zhang, W "A sector-based random routing scheme for protecting the source location privacy in wsns for the Internet of Things.", Future Gener. Comput. Syst. 2019, 96, 438–448. [CrossRef]

[22] Wang, H. Han, G. Zhu, C. Chan, S. Zhang, “A trace cost based source location privacy protection scheme in wsns for smart cities”, Future Gener. Comput. Syst. 2020, 107, 965–974. [CrossRef]

[23] Li, F., et al. “An efficient anonymous communication scheme to protect the privacy of the source node location in the Internet of Things.”, *Security and Communication Networks, 2021*, 6670847. [CrossRef]

[24]Liu, A., et al. “Secure and energy-efficient disjoint multipath routing for WSNs.”, *IEEE Transactions on Vehicular Technology*, Vol 61(7), pp. 3255–3265, 2012. [CrossRef]

[25] Barati, A., Movaghar, A., & Akbari, M “A dynamic and multi-level key management method in wireless sensor networks (WSNs)”, Wireless Personal Communications, 114(2), 1023-1042, 2020. https://doi.org/10.1007/s11277-020-07292-1.

[26] Manjula, R., Koduru, T., & Datta, R. “Protecting source location privacy in IoT-enabled wireless sensor networks: The case of multiple assets”, *IEEE Internet of Things Journal, Vol.9*, pp. 10807–10820,2021. [CrossRef]

[27] Raja, M., & Datta, R, “An enhanced source location privacy protection technique for wireless sensor networks using randomized routes”, *IETE Journal of Research, Vol.64*(6), pp.764–776, 2018. [CrossRef]

[28] Ozturk, C., Zhang, Y., & Trappe, W, “Source-location privacy in energy-constrained sensor network routing”, *Proceedings of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks*, Washington, DC, USA, 25 October 2004, 88–93.

[29] Kamat P., Zhang, Y., Trappe, W., & Ozturk, C “Enhancing source-location privacy in sensor network routing”, *25th IEEE International Conference on Distributed Computing Systems (ICDCS'05)*, Columbus, OH, USA, 6–10 June 2005, 599–608.

[30] Wang, W., Chen, P., & Wang, L “A source-location privacy protocol in WSN based on locational angle”, *Proceedings of the IEEE International Conference on Communications*, Beijing, China, 19–23 May 2008, 1630–1634.

[31] Wang, H., Sheng, B., & Li, Q., “Privacy-aware routing in sensor networks”, *Computer Networks,* (2009), *53*(9), 1512–1529. [CrossRef]

[32] Xi, Y., Schwiebert, L., & Shi, W, “Preserving source location privacy in monitoring-based wireless sensor networks”, *Proceedings of the 20th IEEE International Parallel and Distributed Processing Symposium*, Rhodes Island, Greece, 25–29 April (2006). p.p. 425. [CrossRef]

[33] Zhang, L “A self-adjusting directed random walk approach for enhancing source-location privacy in sensor network routing”, *Proceedings of the 2006 International Conference on Wireless Communications and Mobile Computing*, Amman, Jordan, 17–20 September, 33–38, (2006).

[34] Li, Y., & Ren, J, “Preserving source-location privacy in wireless sensor networks”, *Proceedings of the 2009 6th Annual IEEE Communications Society Conference on Sensor, Mesh, and Ad Hoc Communications and Networks*, Rome, Italy, 22–26 June 2009, pp1–9. [CrossRef]

[35] Li, Y., Ren, J., & Wu J, “Quantitative measurement and design of source-location privacy schemes for wireless sensor networks”, *IEEE Transactions on Parallel and Distributed Systems, 2011, 23*(7), 1302–1311. [CrossRef]

[36] Manjula, R., & Datta, R. (2018), “A novel source location privacy preservation technique to achieve enhanced privacy and network lifetime in WSN”, *Pervasive and Mobile Computing, 44*, 58–73. https://doi.org/10.1016/j.pmcj.2018.01.006