International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2021/209708 Volume 8, Issue 4, July – August (2021)

RESEARCH ARTICLE

Source Location Privacy for Geographical Routing in Wireless Sensor Networks: SLPGR

Manjunath D R
Department of Computer Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India.
manjunathdrcs@gmail.com

Anil Kumar B
Department of Computer Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India.
abnaidu90@gmail.com

Received: 13 July 2021 / Revised: 30 July 2021 / Accepted: 11 August 2021 / Published: 28 August 2021

Abstract – Challenges in the military, environment, medical, industrial, home, traffic applications, and agriculture extend the scope of Wireless Sensor Networks. Data security over wireless networks is a challenge because of the presence of malicious and non-malicious users, whose purpose is to intercept communication or prevent the transmission of data by real users to perform data theft. To improve the location privacy in geographical routing, a hash-based location privacy-preserving scheme and fake source identification in grid-based geographical routing protocol in WSN are presented. In the SLPGR approach, SHA-256 hash encoding is implemented which hides the location information from attackers. The proposed fake source identification guarantees that the fake source and real source nodes are situated on different quadrants and have enough distance between them. The Findings indicate that the SLPGR model's packet delivery ratio is further 278 % enhancement contrast to the tree-based diversionary routing, and more than 38 % compared to the CASER random walking system. The safety duration of the proposed method increases approx. 13% more than the tree-based diversionary routing and 11% more than CASER random walk routing. Energy consumption of the proposed method is lower by 3 times than the tree-based diversionary routing method, 1.4 times lower than CASER random walk routing. The comparative analysis of the SLPGR method shows 3 times lesser delivery miss ratio than tree-based devolutionary routing and 2.6 times lesser delivery miss ratio than CASER routing scheme.

Index Terms – Fake Source, Wireless Sensor Networks, Source Location Privacy, Geographical Routing, SHA-256.

1. INTRODUCTION

Geographic routing protocols in WSNs dealing with sensitive data, especially military and healthcare applications, needs strong security features [1]. Sensor networks are especially exposed to attackers, who may interfere with the proper network operation and thus eliminating out all the potential benefits of this technology. Geographical routing protocols use the geographic location information [2] of nodes provided by GPS (or other systems) in the data forwarding process. Opponents can take advantage of these unique features of WSNs to launch a variety of attacks against the network [3,4].

Source Location Privacy (SLP) [5,6] indicates the ability to protect the source sensor node location that report event data to a destination. The physical sensor node itself is not particularly attractive to invaders. The attacker is anxious to find out the data source because its location is directly related to the location of events. The sensory area and these events may be associated or valued with individuals' resources. Due to the challenging nature of the situation, this issue has attracted the research group attention. Many solutions have been devised to deal with its passive opponents [7], but very few authors focus on the active threats. Consider the panda hunter game [8] shown in figure 1 as an example where nodes detect a panda's location in a region. The node near the panda identifying the panda notifies the base station by transmitting a message through the relaying nodes. To kill the panda, the hunter traces WSN messages to the original node where the panda was perceived. A similar problem like how to protect the panda from a hunter? occurs in other domains also: for example: monitoring a forest fire, tracking grenade's location at the border in the military, monitoring doctors and patients at the hospital [9], and friendly soldiers tracking on the Warfield [10]. To protect the content from the adversary, it is a must to hide the source location of the content. The requirement of SLP is the privacy of messages swapping between nodes. This work focuses on the issue of Source Location Privacy.

The contributions of this manuscript are as follows:

- Put forward a privacy preserving scheme, Source Location Privacy for Geographical Routing (SLPGR) which

RESEARCH ARTICLE

provides two levels of security for SLP against a skilled adversary. In the SLPGR approach, SHA-256 hash encoding is implemented which hides the location information and protects the source location information against ID analysis attackers.

- The proposed a fake source identification Scheme which guarantees that the source node and fake source are situated on different quadrants and have enough distance between them.
- The proposed technique ensures that successfully delivered packets are transmitted through very diverse routes and it is confusing for attackers to backtrack towards the location of the source sensor node.

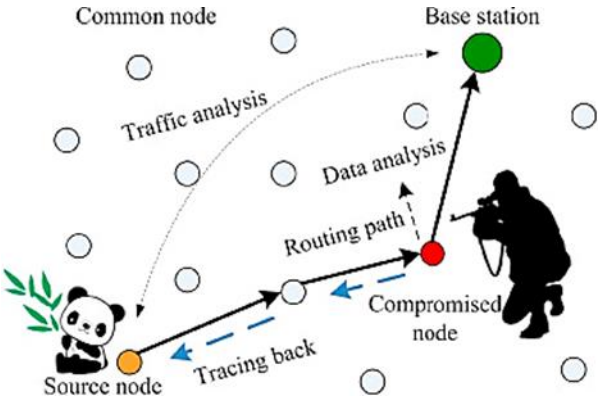


Figure 1 Panda Hunter Game [8]

The rest of this article is organized as follows. Section 2 provides related research work relevant to SLP. Section 3 provides adversary model. The proposed SLPGR Model is overviewed in Section 4, the security analysis is described in Section 5. Lastly, the paper is concluded in Section 6.

2. RELATED WORKS

Most early works deal with only one type of attack but not by a variety of attacks that may start up against routing Protocol. The Geographical-based secure and effective Cost Aware Safe Routing (CASER) protocol has been proposed by Tang.et.al [11] for WSNs to increase network lifetime and balance energy consumption. CASER supports two routing methods for message forwarding. To save energy consumption, nodes message forwards using the shortest path technique. Then, to secure the privacy of the source node location, the random walk technique system is used to make routing unpredictable for jamming prevention and source privacy. Sadly, once the opponents captured a message in a random walk, opponents be able to get zone information or direction available in the message header. A tree-based routing (TR) scheme proposed by Long.et.al.[12]. This scheme resulted in route diversion and there exists a dummy source at every region end of the route diversion to mislead the attacker. The thought is to utilize the non-hotspot areas to boost energy efficiency and also a creation of many paths leads to crossing themselves which creates an issue for an attacker. Multiple routes and crossover additions increase packet path length, and reduces packet delivery ratio. Limitations can further affect the performance of network by provoking a energy hole problem and reducing the lifetime of network [13]. Manivannan.et.al [14] explains different secure privacy protection and authentication techniques and provide an overview of issues related to message delivery, privacy and authentication. It was also observed that the most existing works [11-16] carried out were on source location privacy, designed only for traffic analysis and traceback attacks without considering node id analysis attacks. A hybrid online algorithm is introduced in[17], using directed random walks was introduced for the allocation of duplicate sources to reduce the schemes energy consumption. The Scheme also improves packet delivery ratio by reducing the number of duplicate packets on the network and reduces the number of packet collisions. Three privacy protection for SLP schemes [18], proposed clustering method for one type of privacy protection context. These schemes include Dynamic Shortest Path, Dynamic Tree, and Mixed. Multiple traffic hotspots are created by the clustering method, so that attention can be moved from base station to newly created hot spot. It was also observed that the most existing works carried out were on source location privacy, designed only for traffic analysis and traceback attacks without considering node id analysis attacks. In this work, to prevent an opponent from getting the source node's ID by checking the content of the message, and to provide strong source location privacy, a hash-based location privacy-preserving scheme and fake source identification in grid-based geographical routing protocol [19] in WSN are presented. According to the geographical routing approach, a node transmits packets based on geographical location coordinates, so preserving location privacy is very crucial in geographical routing protocols.

3. ADVERSARY MODEL

The Adversaries, in this work, are assumed by the following characteristics:

Node ID	grid ID	Location coordinates	Neighbouring Node ID	Neighbour Grid ID	Distance	Energy Level of Nodes	Data
Hash	Hash	Hash	Delete/ Update	Delete/ Update	Delete/ Update	Delete/ Update	Encrypt/Decrypt

Table 2 Routing Table and their Corresponding Tasks for Location Privacy

RESEARCH ARTICLE

- Adversaries have insufficient energy resources, adequate computation power, and sufficient memory for storage of data.
- When the event is detected, the immediate sender can be determined by examining the direction and strength of the received signal.
- The attacker does not intervene with the network's proper functioning, as these behaviors can be easily identified. Attackers, however, can conduct eavesdropping on communications.
- The attacker will be able to observe traffic in any particular important area to them and receive all the messages broadcast in that area. Nevertheless, it is thought that opponents cannot monitor the entire sensor network.

4. PROPOSED MODEL

The proposed SLPGR scheme is explained in this section.

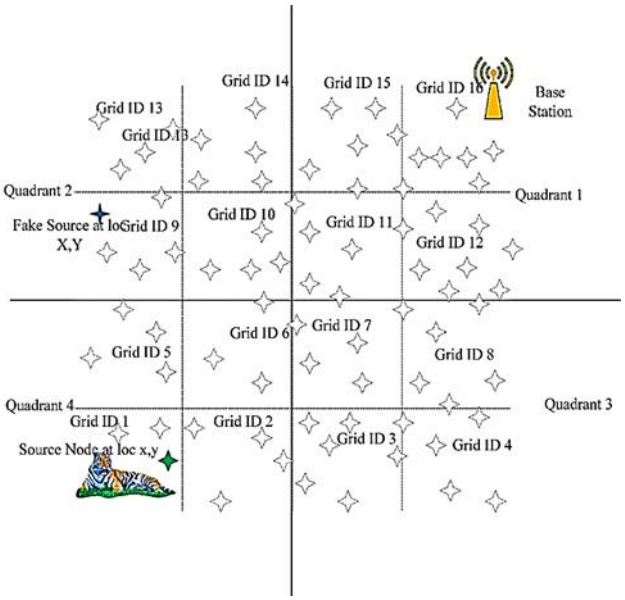


Figure 2 SLPGR Model Schematic

In the SLPGR model, it is assumed that an equal-sized four quadrants square grid and a random distribution of sensor nodes in a non-uniform manner with the static base station at one of the corners in the grid, as manifested in Figure 2. SLPGR method consists of four steps: (i) Network deployment and Grid ID assignment to every sensor nodes by the base station; (ii) Hash Code computation by Base station by applying SHA-256 hash Algorithm on location co-ordinates, Node ID, Grid ID; (iii) Detection of Activity d and Source Node identification; and, (iv) Selection of Grid ID for Fake Source by the Base Station. The complete flow chart for SLPGR source location privacy is presented in figure 3. Division of Quadrants and grid Id initialization is based on Base Station location. In Figure 2 Base Station is located in the top right-most corner, the Quadrants division is from the base station side and the grid id initialization is from the opposite corner side (bottom leftmost corner, in this case), as shown in figure 2.

4.1. Network Model

It is considered that WSNs follow the event-driven data reporting model used for monitoring purposes, such that individual sensor nodes deliver data packets to the base station as soon as the associated phenomena are observed. Thus, after many forwarding hops, all the data is transmitted at the base station. It is assumed that the network is composed of N number of sensor nodes are distributed over a geographical area. Sensor nodes are distributed non-uniformly and randomly in a field, Whose size is W ×H where W denotes the network width and H denotes the height of the network however W and H are considered similar values.

The coordinates of sensor nodes are generated randomly in the considered region as $Coord(x, y) = random(N, 2) \times W$. The complete network region is divided into l×l numbered grids where l represents the grid size, which contains a random number of nodes. Sensor nodes cover a wide area so that an attacker can monitor and control only a small segment of communications at any given moment. Also, it is presumed that each sensor node is aware of its neighbours and that the network connectivity is high. Sensor Nodes allow the next communication hop to be selected from different neighbouring nodes.

Input: Location-coordinates, Node ID, and Grid ID

Output: provides the Hash Code.

1. Initialize eight Hash variables τ_1 to τ_8 with first 32 bits of the square roots fractional part of the first 8 prime numbers' 2 to 19.
2. Initialize sixty-four Round Constants r_1 to r_{64} with first 32 bits of the cube roots fractional part of the first 64 prime numbers, i.e., 2 to 311.
3. Location Information (X, Y Coordinates, Node ID, Grid ID) Input is represented in chunks of 512 bits.
4. For multiple lengths of 512 bits, the input message is supplied by the following: First, append a bit '1', then append m bits '0 '. The input messages length $l < 2^{64}$ is expressed exactly with 64 bits, and are inserted at the message end. Such that l + 1 + m + 64 is a multiple of 512 bits, where l is the length of the initial message bits.
5. For each block, M (512 bits), construct a 64-scheduled array, ω, of 32 bits need to be created and populated.

RESEARCH ARTICLE

6. To populate the array, the message is broken down into 16 32-bit blocks, which are copied to schedule array, W_1 toW_{15}.

7. The remaining 48 elements are gained using the formula

$$\omega_i = \sigma 1(\omega_{i-2}) + \omega i - 7 + \sigma 0(\omega_{i-15}) + \omega_{i-16}, \quad 17 \le i \le 64.$$

Where $\sigma 0 = (\omega_{i-15}\text{ROR } 7) \oplus (\omega_{i-15}\text{ ROR } 18) \oplus (\omega_{i-15}\text{ SHR } 3)$

$\sigma 1(= (\omega_{i-2}\text{ROR } 17) \oplus (\omega_{i-2}\text{ ROR } 19) \oplus (\omega_{i-2}\text{ SHR } 10)$

8. for t=1 to N //compression Function

-do 64 rounds

-Initialize temporary working variables

$$(i, j, k, l, m, n, o, p) = \left(\tau_1^{(t-1)}, \tau_2^{(t-1)}, \tau_3^{(t-1)}, \tau_4^{(t-1)}, \tau_5^{(t-1)}, \tau_6^{(t-1)}, \tau_7^{(t-1)}, \tau_8^{(t-1)}\right)$$

$$p = o, o = n, n = m, m = l + T_1$$

$$l = k, k = j, j = i, i = T_1 + T_2$$

Where $T_1 = p + S_1 + ch + r_l + \omega_i$

$$T_2 = S_2 + maj$$

$$S_1 = (m \text{ ROR } 6) \oplus (m \text{ ROR } 11) \oplus (m \text{ ROR } 25)$$

$$ch = (m \,\&\, n) \oplus (m^{-} \,\&\, o)$$

$$S_2 = (i \text{ ROR } 2) \oplus (i \text{ ROR } 13) \oplus (i \text{ ROR } 22)$$

$$maj = (i \,\&\, j) \oplus (i \,\&\, k) \oplus (j \,\&\, k)$$

- compute the new value of τ_j^t

$$\begin{pmatrix} \tau_1^t = \tau_1^{(t-1)} + a,\ \tau_2^t = \tau_2^{(t-1)} + b, \tau_3^t = \tau_3^t + c, \\ \tau_4^t = \tau_4^{(t-1)} + d, \tau_5^t = \tau_5^{(t-1)} + e, \tau_6^t = \tau_6^{(t-1)} + f, \\ \tau_7^t = \tau_7^{(t-1)} + g, \tau_8^t = \tau_8^{(t-1)} + p \end{pmatrix}$$

9. End for

10. The input message hash is simply the hash variables concatenation after the last block is processed

$$H = \tau_1 || \tau_2 || \tau_3 || \tau_4 || \tau_5 || \tau_6 || \tau_7 || \tau_8$$

Algorithm 1 Hash Computation to Protect Source Location Privacy Model Using SHA-256 Hash Algorithm

4.2. Source Node Location Encoding using Secure Hash Function

To introduce anonymity to the network and grid IDs, hash function operation is applied on x, y location coordinates, Node ID, Grid ID, to hide location information of source node, where each node's corresponding hash ID is used for communication between a source node and destination node. Other parameters shown in table 2 of the routing table such as distance between sensor nodes and energy level of the communicating nodes are updated iteratively during a successful communication and the main information aggregated by the sensor is encrypted and transmitted to the neighbouring node which can be decrypted at the receiver end. A message has integrity when the recipient is assured that the message has not been falsified. The recipient recalculates the hash using the same hash function as the sender and then compares that calculated hash with the hash value attached to the message. If the values are the same, the message hasn't been falsified. The Hash generation of input message that consists of location co-ordinates, Node ID, and Grid Id using SHA-256 is presented in Algorithm 1.

4.3. Registration Phase

Once sensors are deployed in the field, nodes start registering with the base station, as shown in figure 4.

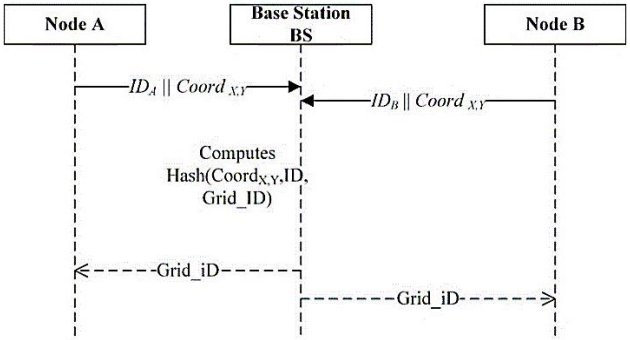


Figure 4 Registration Phase

Step 1: A→ BS: IDA || Coord X, Y; All Nodes Sends its Node-ID and its Location Co-ordinates to Base station.

Step 2: BS → A, B: Grid_ID; Upon receiving Node ID and corresponding location coordinates, Base station computes Grid ID for every node and computes hash code on combining location coordinates, Node ID, Grid ID using SHA-256 hash algorithm and stores this information in base station memory. Base Station sends corresponding Grid ID to every node in the network.

4.4. Packet Transmission Phase

Upon activity detected, Node A becomes a source node, transmits a packet to the base station BS through relay node B, as shown in figure 5.

Step 1. A→B: Hash (CoordX,Y,IDA,Grid_IDA) || BS_CoordX,Y|| EncryptedDATA . As soon as they notice a related phenomenon in their surroundings, Source Node A computes hash code on combining location co-ordinates, Node ID, Grid ID using SHA-256 hash algorithm and sends a packet to neighbor node B based on the proposed geographical routing protocol.

RESEARCH ARTICLE

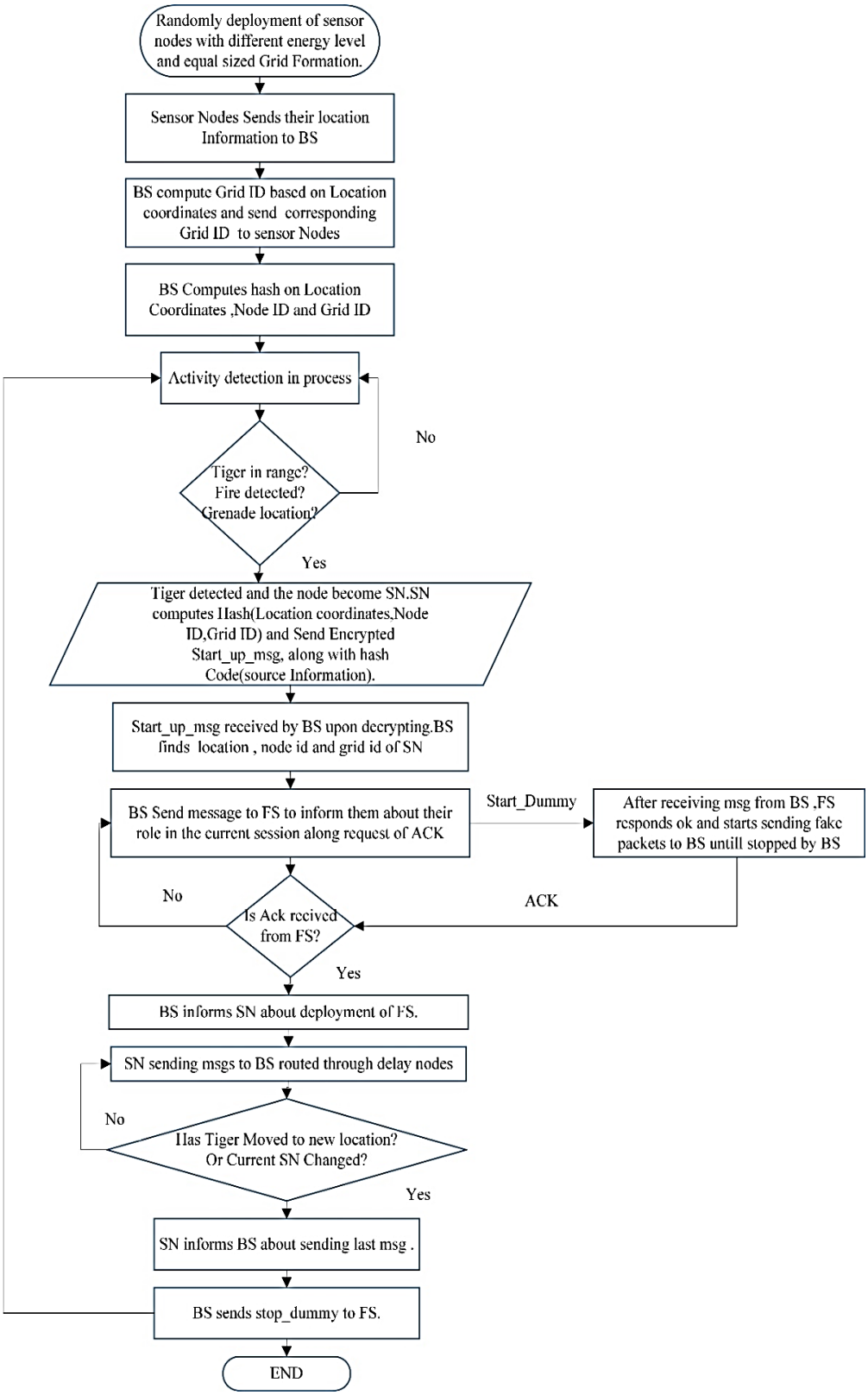


Figure 3 Flow Chart for Proposed Source Location Privacy

RESEARCH ARTICLE

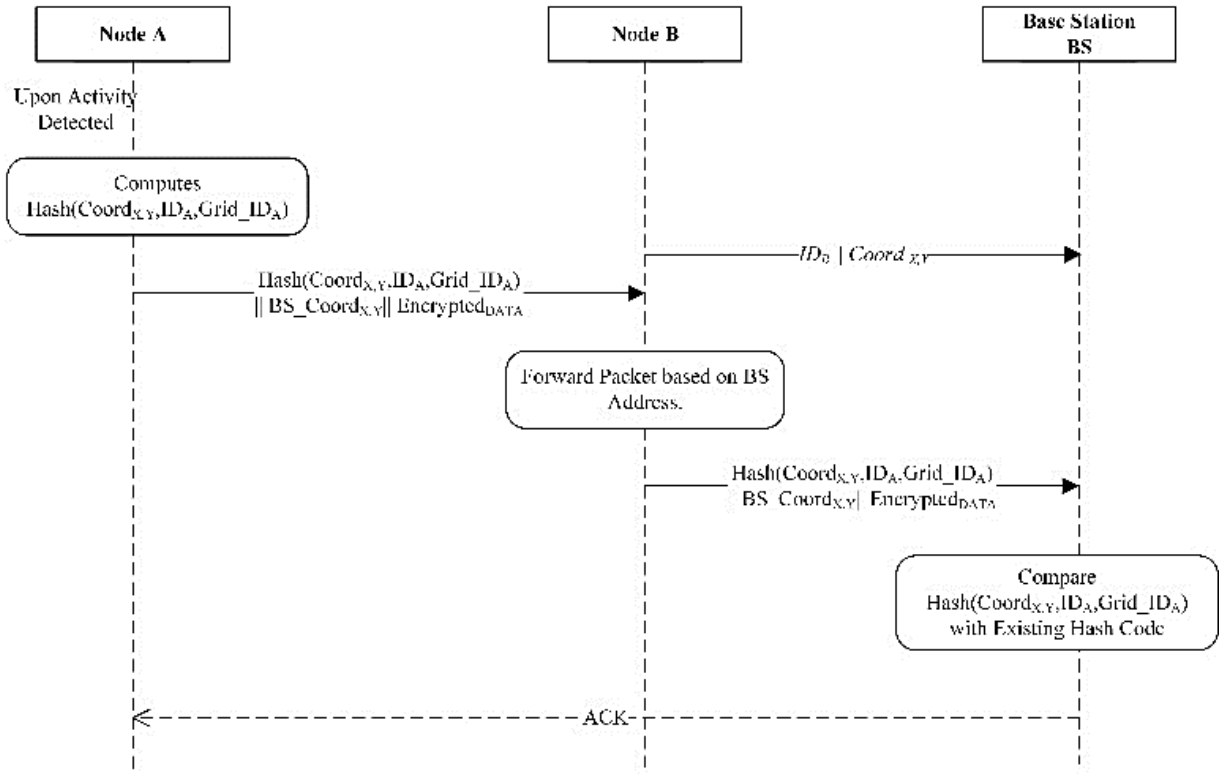


Figure 5 Packet Transmission between Source and Base Station

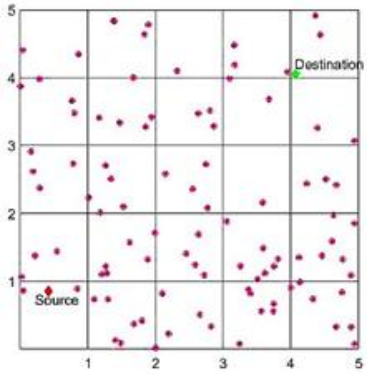
Step 2. B→BS: Hash (CoordX,Y,IDA,Grid_IDA) || BS_CoordX,Y|| EncryptedDATA . Whenever Node B receives data from some source, it forwards packets to its neighbor node or Base station based on Base station coordinates. Each node in the geographical routing protocol is aware of its location, base station location, and neighbours.

Step 3. BS→ A: Ack: Upon received packet from source Node, Base station Decrypt the packet and compare the hash code with the hash codes in its database. If Hash code matches, Base Station identifies the Source Node A and Acknowledges Source Node A. If Hash code doesn't match, then Base station rejects or drop the received packet and it is not acknowledged.

4.5. Fake Source Identification for SLP Model

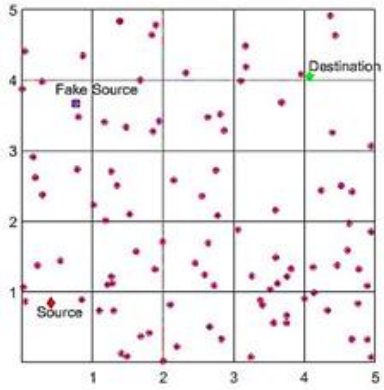
In the proposed model, it is assumed an equal-sized four quadrants square grid and a random non-uniform distribution of sensor nodes with a static base station at one of the corners in the grid, as shown in Figure 2. Division of Quadrants and grid Id initialization is based on Base Station location. In figure 2 Base station is located in the top right-most corner, the Quadrants division is from the base station side and grid id initialization is from the opposite corner side (bottom leftmost corner, in this case), as shown in figure 2. If the activity is detected (ex: a tiger is detected in the grid 1 region) then the proposed Algorithm for identifying grid id to choose fake source is presented in algorithm 2.

Whenever an event is detected, one of the nodes is chosen as the source node in that grid region based on the geographical routing [19]. Source Node generates a single startup message and sent the message along with its hash code to the base station. Now no need to think about detecting the source location, since it is impossible for an attacker to do backtracking on a single message and also impossible to find source location in a piece of packet information since source location is presented in the form of hash code, an only base station is capable of identifying the source node.



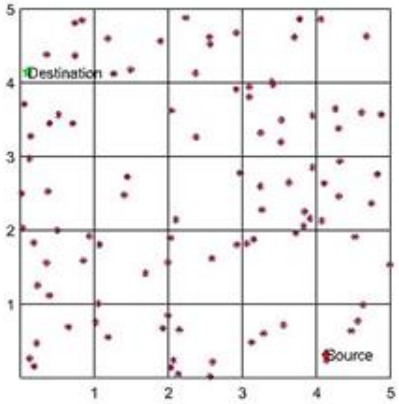
(a)

RESEARCH ARTICLE

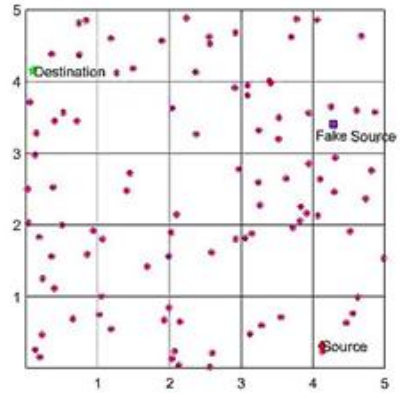


(b)

Figure 6 Fake Source Identification a) Source and Destination at Grid {1,1} and Grid{5,5} Respectively b) Fake Source Identified for Source and Destination at Grid{1,} and Grid{5,5} Respectively

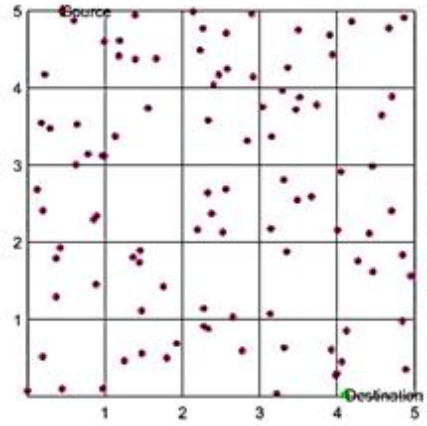


(a)

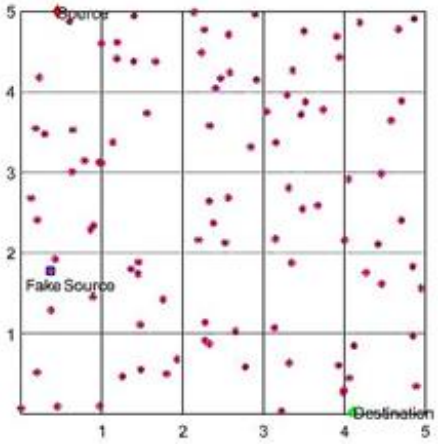


(b)

Figure 7 Fake Source Identification a) Source and Destination at Grid {1,5} and Grid {5,1} Respectively b) Fake Source Identified for Source and Destination at Grid {1,5} and Grid {5,1} Respectively

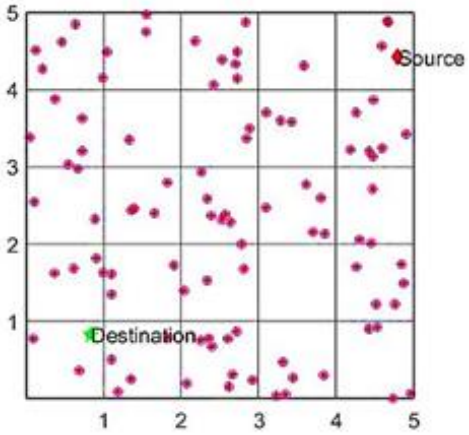


(a)



(b)

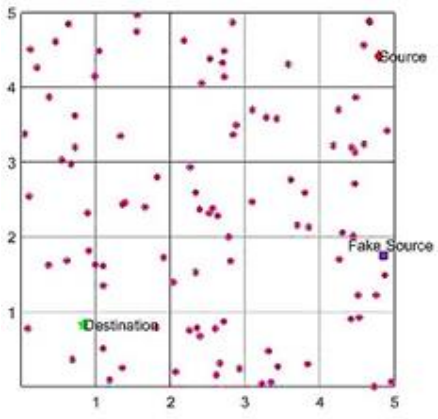
Figure 8 Fake Source Identification a) Source and Destination at Grid {5,1} and Grid {1,5} Respectively b) Fake Source Identified for Source and Destination at Grid {5,1} and Grid {1,5} Respectively



(a)

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2021/209708 Volume 8, Issue 4, July – August (2021)

RESEARCH ARTICLE



(b)

Figure 9 Fake Source Identification a) Source and Destination at Grid {5,5} and Grid {1,1} Respectively b) Fake Source Identified for Source and Destination at Grid {5,5} and Grid {1,1} Respectively

The selection process of a Fake source in a grid region other than the source node includes an algorithm to achieve a sufficient gap between the sources.

After the activity is detected in the grid id 1(Quadrant 4 in this case), the base station selects the grid id having enough distances between the source node grid id (grid id 9 in Quadrant 2 in this case), as shown in figure 2. Fake source grid identification and fake source selection with different locations of source and base station based on a fake source identification algorithm are shown in Figures 6,7,8, and 9.

5. RESULTS AND DISCUSSIONS

In this section, a complete experimental study is presented using PBGR [19] protocol. The SLPGR approach is simulated using MATLAB. The complete experimental parameters are shown in Table 3. In the current simulation scenario, it is considered $l = 5$ hence total of 25 smaller grids are generated which contains multiple nodes or no node in the random distribution. The destination node's coordinates are given as [4.06799, 4.09346]. Now Source Node (Node ID 19) sent a start-up message, corresponding hash address to the Base Station. Upon the first message received, now base station chooses a fake source grid id and select a highest energy sensor node in that grid as the fake source.

Simulation Parameter	Considered Value
Network area (W x H)	1500*1500 m2
The initial energy of nodes E_{init}	1 joule
Total number of nodes	100
Transmission range d_0	50 m
Packet Size	8 bytes
Electronics energy E_{elec}	50nJ/bit
Free space model amplification energy (ε_{fs})	$100 * 10^{-12}$ joules
Multipath model amplification energy ε_{amp}	$0.0015 * 10^{-12}$ joules
Energy threshold ($Energy_{threshold}$	$E_{init}/4$ joules
Data rate (kbps)	2,4,6,8

Table 3 Simulation Parameters

Input: Current grid ID and grid size (l)// current grid id is obtained from (1)

Output: Provides the Fake source grid ID.

1. Provide the grid size and current grid ID as input.
2. Initialize the $Grid_{middle} = \frac{l}{2}$, $Grid_{threshold} = grid_{middle} * l$.

// Upper grids estimation

3. if $(gid + Grid_{threshold}) > l^2$
4. if$(mod(gid, l) < Grid_{middle})$
5. if$(gid > Grid_{threshold})$
6. Then $FS_{grid} = gid - Grid_{threshold} + 1$ end if
7. Else
8. $FS_{grid} = Grid_{threshold} - gid + 1$ end if
9. end if
10. if$(mod(gid, l) > Grid_{middle})$
11. if$(gid > Grid_{threshold})$
12. Then $FS_{grid} = gid - l + 1$ end if //It can be avoided by using fake Source here since all nodes are nearest to Base station Else
13. $FS_{grid} = gid + l - 1$ end if
14. end if
15. end if

//Lower grids estimation

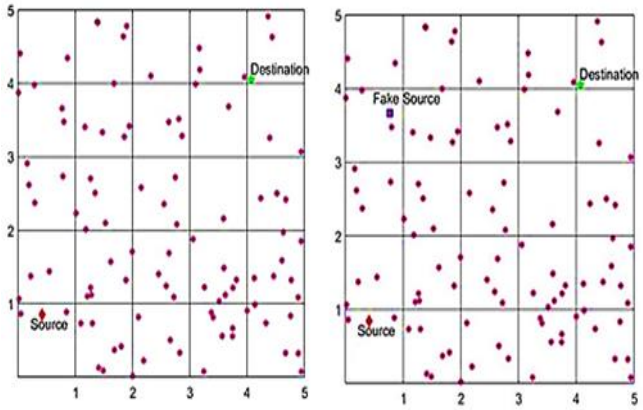
16. if $(gid + Grid_{threshold}) \leq l^2$
17. if$(mod(gid, l) \leq Grid_{middle})$
18. Then $FS_{grid} = gid + Grid_{threshold}$ end if

RESEARCH ARTICLE

19. Else
20. $\text{FS}_{\text{grid}} = \text{gid} + \text{Grid}_{\text{threshold}} - \text{Grid}_{\text{middle}}$ end if
21. end if
22. end if

Algorithm 2 Fake Source Grid Identification Process

In our case, source node grid id is 1, so Source Node belongs to Quadrant 1, According to the SLPGR algorithm, fake source grid identification is as follows, $Grid_{middle} = \frac{5}{2}$, is 3 and $Grid_{threshold} = 15$, so the Fake grid is in the upper level (Quadrant 2). the final grid id is $FS_{grid} = gid + Grid_{threshold}$, so $FS_{grid} = 1 + 15$ =16. Fake source grid identification and fake source selection with different locations of source and base station based on fake source identification algorithm are shown in figure 10. In our case, the base station chooses one of the nodes (highest energy node) in grid 16 as a fake source and sends a message to Fake Source to inform them about their role in the current session along with the request of Acknowledgement. After receiving a message from Base Station, the Fake source responds ACK and commence transmitting fake packets to Base Station until stopped by Base Station. Base Station informs Source Node about the deployment of Fake Source. Source Node now starts transmitting packets to Base Station routed through relay nodes. The intermediate nodes act as relay nodes; they transmit packets to the destination. Relay Nodes do not know which node-initiated source message, since the source address is in the form of hash code, only the base station is capable of identifying the source node.



(a) (b)

Figure 10 Sensor Node Deployment: (a) Identified Source and Destination (b) Identified the Fake Source

5.1. Energy Consumption Model

For implementing the SLPGR protocol, a node S_i to transmit $k - bits$ of information over a distance$dist$, it needs $E_{Tx\_elec}^{(i)}$ energy as shown in Equation 1.1.

$$E_{Tx\_elec}^{(i)}(k, dist) = \begin{cases} k * En_{elec} + k * \varepsilon_{fs} * dist^2, & \text{if } dist < dist_0 \\ k * En_{elec} + k * \varepsilon_{amp} * dist^4 & \text{if } dist > dist_0 \end{cases} \quad (1.1)$$

where, En_{elec} represents the small amount of energy utilized due to underlying hardware. En_{elec} is assumed to be fixed for certain hardware types. Energy utilization due to multi-path amplification is represented as ε_{amp} and the free space channel is represented ε_{fs}, respectively. The transmission distance threshold is indicated by $dist_0$. Distance threshold can be calculated as, $dist_0 = \sqrt{\varepsilon_{fs} / \varepsilon_{amp}}$. Similarly, for a S_i to receive $k - bits$ of information, it needs $E_{Rx}^{(i)}$ energy, as shown in equation 1.2.

$$E_{Rx\_elec}^{(i)} = k * En_{elec} \quad (1.2)$$

where, $dist(x, y)$ represents the Euclidean distance between the source sensor node S_{SN} and it's neighbor sensor node S_{NN} . If the co-ordinate position of S_{SN} and S_{NN} is assumed to be $(\alpha_{\text{x}}, \beta_{\text{x}})$ and $(\alpha_{\text{y}}, \beta_{\text{y}})$ respectively, then Euclidean distance between S_{SN} and S_{NN} is calculated as per Equation 1.3.

$$dist(x, y) = \sqrt{(\alpha_{\text{x}} - \alpha_{\text{y}})^2 + (\beta_{\text{x}} - \beta_{\text{y}})^2} \quad (1.3)$$

The overall energy consumed by a particular path between the source sensor node and the base station node is given by

$$E_{total_{en}} = \sum_{i=1}^{n} Et_i \quad (1.4)$$

Where Et_i *is* the total energy consumed from a specific path between source and destination and 'n' is the total number of hops in a specific path to the destination.

$$Et_i = E_{Rx\_elec}^{(i)} + E_{Tx\_elec}^{(i)} \quad (1.5)$$

5.2. Performance Evaluation

Secure hash encoding of the source node location information is considered the touchstone in our simulations. It is observed that random walk; phantom routing is often used in the related study for comparison for source location privacy. Phantom routing is best in packet delivery ratio and energy efficiency, but it is not capable of providing optimal source location privacy. Whereas random walking provides the best good source location privacy, but it is best in the packet delivery ratio.

For the hash computation of location info (locx,locy,nodeid,gridid) i,e (0.44561.77563,7,1) and the generated hash code is and the computation time for different Hash algorithms are shown in table 4.

In the proposed model, we have two main sources: source node and fake source. Data from different sources are heap up toward the base station. The adversaries usually lives near the

RESEARCH ARTICLE

Base Station at the beginning and then uses backtracking to find the source node.

Hash Algorithm	Hash Code	CPU Time
MD5	63fad8568e6508fba92acb1adfad50e0	25 ms
SHA-1	4b229358cef3f7c43d114c5f634599d5e3094d58	24 ms
SHA-256	1574b8e8966b4c626d2b75d976318207ecf81ec24ce6a1cfaa7d843beb9f4e45	25 ms
SHA-512	daacb289dd737f7abbb21490ebe560f3df7aacc7ac97cf0a423f6400b47cc5439e6d01d696f18921e2783dc7c0d4e334b7d3ecd47fbf7076fc7a9e3cbce93ff1	44 ms

Table 4 Comparison of Different Hash Algorithms

5.2.1. Packet Delivery Ratio

On the other hand, a comparison of the performance of the SLPGR approach is done by considering all security parameters and compare the obtained performance with the existing approach in terms of packet delivery ratio for discrete simulation time. The packet delivery ratio measures the percentage of the total number of packets received successfully at the base station. To compare the performance, a recent work proposed by Tang et al. [11] is considered, known as CASER and tree-based diversionary routing [20]. In this work, the authors presented control parameters that can be adjusted according to the required performance as energy management or security.

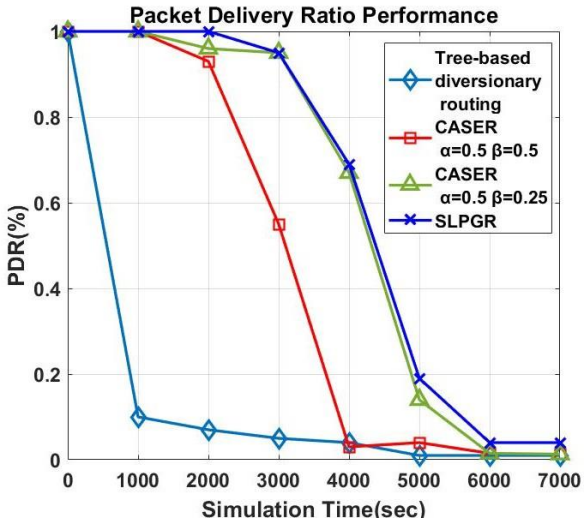


Figure 11 Packet Delivery Ratio Performance

Figure 11 compares the SLPGR scheme packet delivery ratio with tree-based diversionary routing, CASER with window parameters α=0.5, β=0.5 and CASER with parameter window α=0.5, β=0.2515. In tree-based diversionary routing the findings indicate a rapid reduction of the delivery ratio of packets, followed by the CASER protocol, and proposed model. As the SLPGR model is simple, with geographical energy-aware routing paths, so the performance of delivery ratio of packet is better as compared to CASER random walking and tree-based diversionary routing schemes, which outcomes in long routing paths. These extra-long routing paths in CASER and tree-based diversionary routing also introduce delivery latency.

The findings indicate that the SLPGR model's packet delivery ratio is an enhancement of 278 % contrast to the tree-based diversionary routing, and more than 38 % compared to the caser random walking system in case of parameters α=0.5, β=0.5. Since tree-based diversionary routing has fake nodes with a plenty of diversionary long routes, which results to packets collision, so the delivery ratio of packets is worse contrasted to the SLPGR method. CASER also introduce delivery latency because of the extra-long paths. The primary reason for this successful packet delivery ratio is that the fake node and source node reside in separate division, reducing the risk of intersecting paths and thus lower the likelihood of packet crashes. The findings show that the SLPGR model provides higher throughput when compared with the random walk scheme and tree-based diversionary routing while offering effective privacy for the source node location.

5.2.2. Safety Period

SLPGR system demonstrates good results in the safety period when compared to the random walk scheme. The safety periods are described as the time needed to backtrack and capturing the event for an attacker, or it is the maximal time the tiger stays at the source sensor node before it moves to a different location. Figure 12 compares the SLPGR scheme safety period with CASER with window parameters α=0.5, β=0.5 and CASER with parameter window α=0.5, β=0.2, and tree-based diversionary routing, the findings show that the safety duration approximately increases enormously in the proposed model approx. 13% more than the tree-based diversionary routing and 11% more than CASER random walk routing. This is because the fakes node and real source node reside in separate division, so real packets and fake dummy packets come from different sides. The main input in the safety period is the hash encoding, then the fake source. As the attacker near the Base Station attacker receives packets on each side, the backtracking packet is very confusing since the source node and fake source are on different quadrants. Even if an attacker chooses to track someone, he could end up with a fake source, increasing security. In addition, if the attacker analyses the packet, he could not find source location

RESEARCH ARTICLE

coordinates since location information is encoded with hash code, only base station aware of the fake source and source node location, which leads to increasing source location privacy.

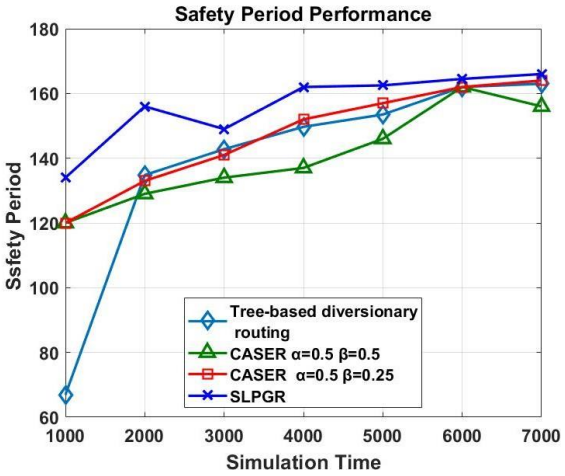


Figure 12 Safety Period Performance

5.2.3. Energy Consumption

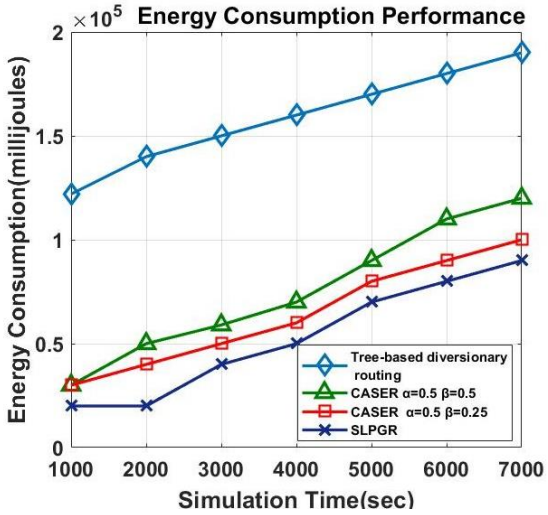


Figure 13 Energy Consumption Performance

Sensor nodes energy consumption depends on the distance between the source sensor node and the sink. Consumption of energy is the total power consumed by the sensor nodes over network lifetime. Optimal outcomes can be accomplished by shortest path routing for power consumption, but the Source Location Privacy is compromised. To have a decent security period, energy consumption must be compromised. This tradeoff is acceptable because the security period must be improved and confusion for the intruder must be increased. In such applications, the rise in energy consumption is worthy of the monitoring of precious resources and safety. The system ensures that the fake source node grid will always be an indifferent quadrant of the source node. The fake source model ensures that whenever the source node is in quadrant 4, the fake source grid will be in quadrant 2 as shown in figure 2.

Assume, if the source node is on quadrant 2, then the eligible fake source grid is on either quadrant 4 or quadrant 3. Suppose if the source sensor node is present in quadrant 3, then the fake node grid is in quadrant 2. Finally, if the source sensor node is situated in quadrant 1, picking a fake node grid can be avoided since the source nodes are close to the base station.

The CASER random walk and tree-based diversionary routing model operates numerous diversionary routes that increase the overall packet's route length, thereby increasing energy consumption. SLPGR model needs energy for each source node for the computation of source node hash data. Source node needs to calculate hash code only once, upon activity detected and routes packets using PBGR [19] protocol, thus lowering consumption of energy. Figure 13 compares the consumption of energy performance of the proposed scheme with tree-based diversionary routing, CASER [11] with window parameters α=0.5, β=0.5 and CASER with parameter window α=0.5, β=0.25. Energy consumption of the proposed method is lower by 3 times than tree-based diversionary routing[10] method, 1.4 times lower than CASER random walk routing with window parameters α=0.5, β=0.5 and 1.2 times lower than CASER random with window parameters α=0.5, β=0.25.

5.2.4. Delivery Miss Rate

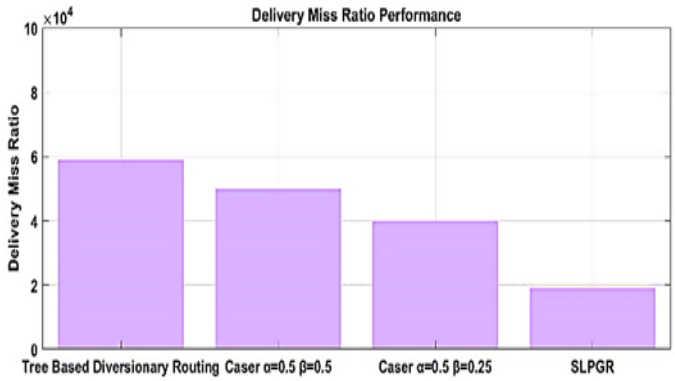


Figure 14 Delivery Miss Rate

The delivery miss rate of the three methods is shown in figure 14. The delivery miss rate reveals the proportion of the number of packets transmitted through the source sensor node but not acquired in the sink node. Due to many distraction routes in the CASER [11] and tree-based scheme [20] leads to the highest proportion of packet loss. These routes run across each other, resulting in a collision of packets, which increases

RESEARCH ARTICLE

the delivery miss rate. The findings shows that the SLPGR method shows 3 times lower delivery miss rate than tree-based devolutionary routing and 2.6, 2 times lesser delivery miss ratio than CASER routing scheme.

5.2.5. Packet Delay Ratio

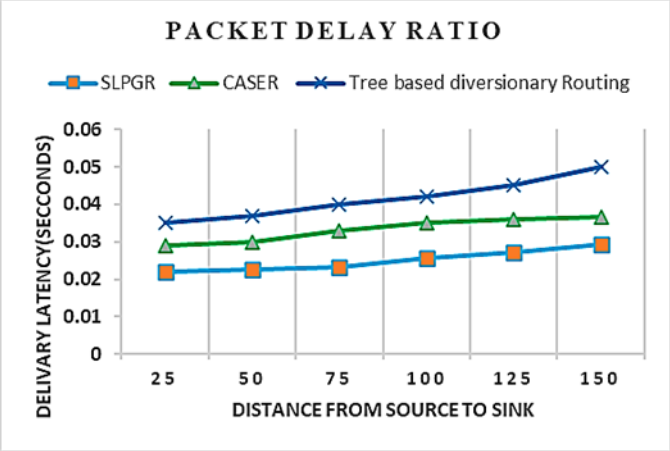


Figure 15 Packet Delay Ratio

Figure 15 shows the latency of proposed SLPGR, Tree based diversionary Routing and CASER under different number of nodes. The proposed SLPGR scheme is more economical than the other schemes. The tree based diversionary routing and random walking in CASER routing scheme has the highest packet delay ratio since it employs long routing paths to transmit the packets. Since the packets are randomly transmitted through relay sensor nodes, so as the number of nodes increases, we can see that the latency also increases.

5.3. Security Analysis

Here it is described some of the key aspects of security analysis results, including source location privacy, ID analysis attacks, and passive attack analysis. To prevent an adversary from getting the source sensor node's ID by checking the matter of the message, the SHA-256 hashing scheme is used to hide the details of source location coordinates, node id, and grid id. To introduce anonymity to the network and grid ids, the base station applies the hash function on every node's location coordinates, node ID's, and grid id's during the network deployment stage, where each node's corresponding hash id is used to establish communication between source and base station so that attackers cannot add duplicate packets to the network or act as a legitimate sensor node; they do not modify the information of the message. More importantly, attackers cannot create the ID and information of the source node's location by analyzing the contents of the message or the network in the SLPGR protocol. In this field of research, it is more important to be aware of securing the source node location of the, because compromising it can outcome in the reduction of assets value for those observing the network. As the attacker near Base Station receives packets on each side, the backtracking packet is very confusing since the source node and fake source are on different quadrants. Even if an attacker chooses to track someone, he could end up with a fake source, increasing security.

6. CONCLUSION

Source location privacy (SLP) for a privacy-crucial WSN applications, such as tracking and monitoring is becoming a key feature. In this work, a Security scheme for SLP is presented that contributes two levels of security. Firstly, the identities and location information of source nodes are encoded using the hash function, only the base station is capable of identifying the source by comparing hash code with stored hash codes in memory. Even if attacker performs packet analysis, they cannot find node id, as well as location information of source node. Secondly, a novel method is proposed for identifying a fake source in a randomly distributed, grid-based geographical energy-aware routing protocol, which guarantees that the source node and fake source are always on different quadrants and have enough long distance between them. This ensures that traffic from separate parts of the network is moving towards sink node to prevent attackers from backtracking the source node by traffic analysis. The extensive simulations performed have shown that the devised SLPGR method is capable of providing a solid level of privacy protection with an average consumption of energy very close to optimal. The results show that the SLPGR approach has more packet delivery ratio, more safety period performance, less energy consumption, and lower delivery miss rate as compared to CASER and tree-based diversionary routing scheme. In the SLPGR approach, SHA-256 hash encoding is implemented which hides the location information from attackers. In the future it is required to identify an energy-efficient, lightweight secured hash algorithm to encode source location information. Additionally, in further work, it is needed to study the implementation of flooding ring of sensor nodes near to base station to further confuse attackers of the backtracking source node.

REFERENCES

[1] Naghibi, Maryam and H. Barati. "EGRPM: Energy efficient geographic routing protocol based on mobile sink in wireless sensor networks." Sustain. Comput. Informatics Syst. 25 (2020): 100377.
[2] Ahmad Raza Hameed, Saif ul Islam, Mohsin Raza, Hasan Ali Khattak, Towards Energy and Performance-aware Geographic Routing for IoT-enabled Sensor Networks, Computers & Electrical Engineering, Volume 85,2020,106643,
[3] Z. Qian, Q. Xiaolin, and D. Youwei, "Intelligent silent zone for source-location privacy based on context-awareness in WSNs," Transactions of Nanjing University of Aeronautics and Astronautics, vol. 35, no. 1, pp. 203–218, 2018.
[4] C. Gu, M. Bradbury, J. Kirton, and A. Jhumka, "A decision theoretic framework for selecting source location privacy aware routing protocols in wireless sensor networks," Future Generation Computer Systems, vol. 87, pp. 514–526, 2018.

RESEARCH ARTICLE

[5] H. Wang, G. Han, W. Zhang, M. Guizani and S. Chan, "A Probabilistic Source Location Privacy Protection Scheme in Wireless Sensor Networks," in IEEE Transactions on Vehicular Technology, 68, 6, pp. 5917-5927, June 2019..
[6] R. Rios, and J. Lopez, "Analysis of Location Privacy Solutions in Wireless Sensor Networks", IET Communications, vol. 5, pp. 2518 - 2532, 2014.
[7] Mutalemwa, L.C.; Shin, S. Secure Routing Protocols for Source Node Privacy Protection in Multi-Hop Communication Wireless Networks. Energies 2020, 13, 292. https://doi.org/10.3390/en13020292
[8] Jinfang Jiang, Guangjie Han, Hao Wang, Mohsen Guizani, A survey on location privacy protection in Wireless Sensor Networks, Journal of Network and Computer Applications, Volume 125,2019, Pages 93-114. https://doi.org/10.1016/j.jnca.2018.10.008
[9] Nidhi Sharma, Ravindra Bhatt, Privacy Preservation in WSN for Healthcare Application, Procedia Computer Science, Volume 132, 2018, Pages 1243-1252, ISSN 1877-0509. https://doi.org/10.1016/j.procs.2018.05.040
[10] L.C. Mutalemwa, S. Shin Achieving source location privacy protection in monitoring wireless sensor networks through proxy node routing Sensors, 19 (2019), p. 1037, https://doi.org/10.3390/s19051037
[11] Tang, D., Li, T., Ren, J., & Wu, J. Cost-Aware Secure Routing (CASER) Protocol Design for Wireless Sensor Networks. IEEE Transactions on Parallel and Distributed Systems, 26, 960-973, 2015. https://doi.org/10.1109/TPDS.2014.2318296
[12] 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 2014, 2, 633–651. https://doi.org/10.1109/ACCESS.2014.2332817
[13] N. Jan and S. Khan, Energy-Efficient Source Location Privacy Protection for Network Lifetime Maximization Against Local Eavesdropper in Wireless Sensor Network (EeSP). Hoboken, NJ, USA: Wiley, Aug. 2019
[14] Manivannan D., Moni S.S., Zeadally S. Secure authentication and privacy-preserving techniques in Vehicular Ad-hoc NETworks (VANETs). Veh. Commun., 25 (2020)
[15] L. Lightfoot, Y Li, J. Ren. STaR: design and quantitative measurement of source-location privacy for wireless sensor networks. Security and Communication Networks. 2016, 9(3): 220-228. https://doi.org/10.1002/sec.527
[16] Mutalemwa, L.C.; Shin, S. Strategic Location-Based Random Routing for Source Location Privacy in Wireless Sensor Networks. Sensors 2018, 18, 2291. https://doi.org/10.3390/s18072291
[17] Bradbury, M.; Jhumka, A.; Leeke, M. Hybrid online protocols for source location privacy in wireless sensor networks. J. Parallel Distrib. Comput. 2018, 115, 67–81.
[18] Al-Mistarihi MF, Tanash IM, Yaseen FS et al (2020) Protecting source location privacy in a clustered wireless sensor networks against local eavesdroppers. Mob NetwAppl 25:42–54
[19] D. R. Manjunath and S. N. Thimmaraju, "A blind path geographical energy aware routing protocol for wireless sensor networks," 2019 1st International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE), Bangalore, India, 2019, pp. 196-200, doi: 10.1109/ICATIECE45860.2019.9063840.
[20] Jan N, Al-Bayatti AH, Alalwan N, Alzahrani AI. An Enhanced Source Location Privacy based on Data Dissemination in Wireless Sensor Networks (DeLP). Sensors (Basel). 2019;19(9):2050.doi:10.3390/s19092050. https://doi.org/10.3390/s19092050.

Authors

Dr. Manjunath D R is currently working as Assistant Professor in Dept. of CSE , Dayananda Sagar Academy of Technology and Management, Bangalore, received B.E,M.Tech and Ph.D. in Computer Science and Engineering from Visvesvaraya Technological University, Belagavi, Karnataka, India.

Mr.Anil Kumar B is currently working as Assistant Professor in Dept. of CSE , Dayananda Sagar Academy of Technology and Management, Bangalore, received B.E and M.Tech in Computer Science and Engineering from Visvesvaraya Technological University, Belagavi, Karnataka, India, He is currently pursuing the Ph.D. in Computer Science and Engineering from Visvesvaraya Technological University.

How to cite this article:

Manjunath D R, Anil Kumar B, "Source Location Privacy for Geographical Routing in Wireless Sensor Networks: SLPGR", International Journal of Computer Networks and Applications (IJCNA), 8(4), PP: 422-434, 2021, DOI: 10.22247/ijcna/2021/209708.