# Techniques to Preserve Source Location Privacy in Wireless Sensor Networks

**C N Chinnaswamy**
Associate Professor, Dept. of ISE
National Institute of Engineering, Mysuru

**Rajat R Hegde, Rajendra Prasad S, Rakesh V Maradi and Vijay**
Information Science and Engineering,
National Institute of Engineering, Mysuru

## Abstract

*Wireless Sensor Network (WSN) has made it's impact on various fields like Military, Wildlife protection and many other . In WSN sensor nodes are deployed over a region , nodes which senses the changes within it's range and sends the information to sink . We need to provide privacy to source node from intruder , who may trace the location of source node and steal more information .While monitoring the target, the revealed information about the subject can be misused by the adversary. Hence, our aim is to hide the source location from the adversary. It is quite difficult to efficiently achieve the source location privacy although the confidentiality of the messages can be well assured through data encryption. In WSN, the source location privacy is more complex due to the fact that the nodes comprise energy efficient and less computational and less storageable devices. In this paper, we provide various techniques like Random Walk, Multiple Phantom Node, Base Line Flooding, STaR routing to preserve the location of source node from intruder.*

*Index Terms- Dynamic Routing, Adversary, Wireless Sensor Networks, Context Privacy, Source Location Privacy, Phantom Node, Random Walk, WSN privacy, Baseline Flooding, STaR Routing.*

## I. INTRODUCTION

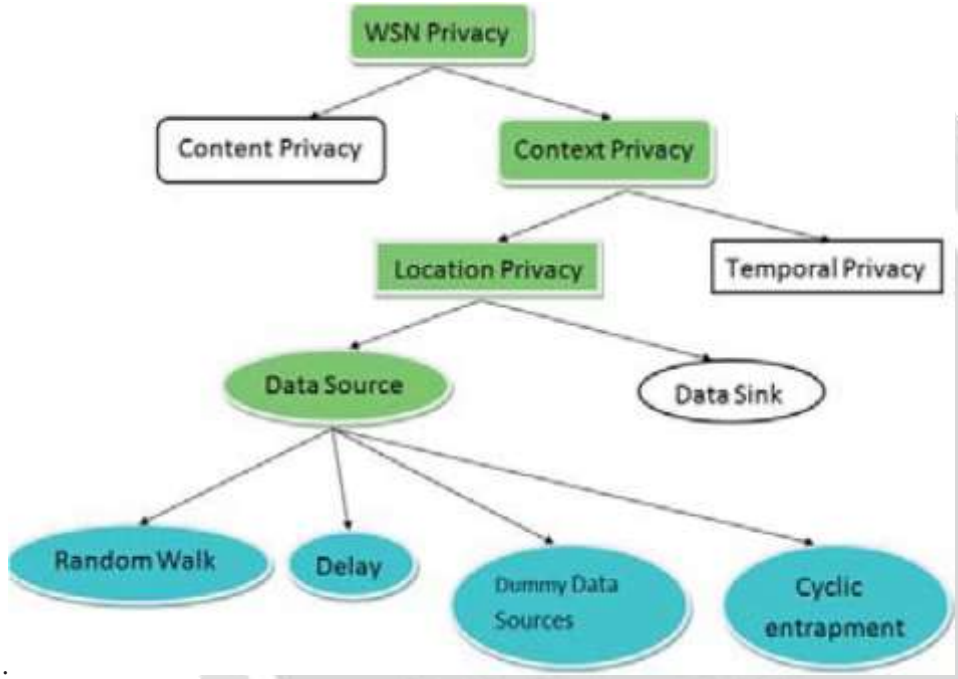
In today's world, the advancement in wireless communications has enabled the development of sensor-based networks. In comparison to the traditional networks, the WSN has become a new adopted network structure. WSN inherently based on the wireless communications, which is basically an open media. The wireless communications are more prone to privacy and security threats than the wired one. In wireless domain, anybody equipped with an sufficient hardware can intercept and monitor the wireless network communication. An adversary may use high frequency radio transceivers to monitor the network communications from a distance. It is very likely that the source location can be easily identified by the adversary through tracing its messages. One of the major applications of sensor-based network in privacy domain is subject monitoring and tracking. The sensor nodes are randomly deployed in a sensing region to inspect its object of interest, which is called as a subject.

## II. OVERVIEW ON SOURCE LOCATION PRIVACY

Wireless sensor networks are composed of many small sensor nodes that can sense, collect and spread information for different types of applications. Sensing of data includes sensing physical quantity such as temperature, humidity, pressure, radiation etc. Wireless sensor nodes have limited storage, computing power, and energy supply. The Wireless Sensor Network (WSN) is technology in which Sensor nodes are dispersed over a certain region, these sensor nodes forms the small network and nodes communicate based on Wireless Communication. Sensor nodes which senses the some particular object that are present within it's range and this process of sensing the objects is called event. After the deployment of sensor nodes, the nodes are left unattended for most of the practical applications. One of the major application is subject tracking and monitoring where not only

data but also the location of the sensor node needs to be preserved . Privacy can be defined as "a state in which one is not observed or disturbed by other people". It is the state of being free from public attention. Privacy in wireless sensor networks includes hiding of nodes location, confidentiality, availability and integrity of messages etc. It can be broadly classified into two parts: Content Privacy and Context Privacy. Content privacy deals with the protection of data that is being communicated between sensor nodes while context privacy deals with the context related to the information such as source location, destination (sink) location and time at which the message was created. Context privacy includes hiding the identity and the locality of each node, and hiding the flow of traffic among the nodes. Our focus is on Source Location Privacy(SLP) as shown in **fig1**

**Fig. 1: Privacy in WSN**




.

We first introduce the concept of SLP on the basis of example of panda-hunter game. The main idea behind the provided example is that the hunter only analyzes the routing pattern of the messages transmitted to locate the panda. Here the Nodes are scattered in an area to trace the location of panda. When a node senses the existence of panda within its range it informs the sink by sending messages through intermediate nodes. In the interim, if the hunter also listens to the message arrived from the source while moving in the network. The hunter starts tracing the message to locate the source node in order to kill the panda. So, the question arises over here is that “how do we protect the panda from hunter?” To protect it from the adversary, we must hide its location i.e. preserving source location privacy (SLP). Hiding the source location is a challenging task as there are several aspects that affect the efficiency of the solution as the mobility of the nodes. Another influencing factor is the need of protecting the source from the adversary. An adversary that compromises the node has different capabilities than those that cannot compromise. The capability of an adversary to view the traffic within the network is also an important issue and it may differ. Lets get the outlook of panda hunter game in **Fig 2.**

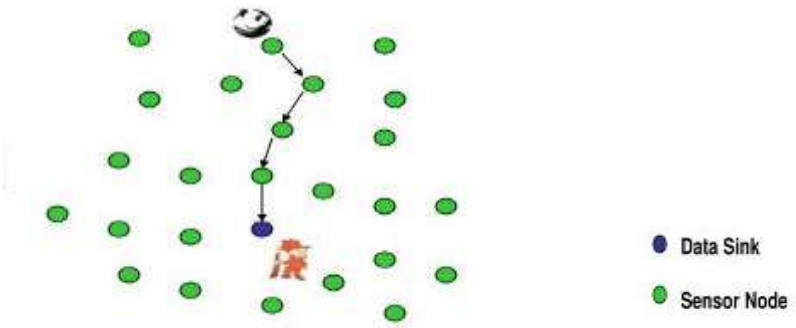



**Fig. 2: Panda Hunter Game**

## III. ADVERSARY

This section provides a summary of the adversary capabilities considered in numerous threat models. The threat models are classified on the basis of four disparate properties: the adversary behavior, its network view, the resources it has and the information exposed.

### *A. Adversary behavior*

The behavior of the adversary is classified on three distinctive criteria:

 Accessibility: The adversary can be external or internal. An internal adversary can alter the functionality of a node in the network. But an external adversary cannot perform this.

 Interference: The adversary can be passive or active. A passive adversary does not interfere in the proper functioning of the solution whereas an active does alter the behavior of the nodes.

 Compliance: The adversary can be dishonest or semi-honest. A dishonest adversary does not abide by the protocols in the network whereas the semi-honest adversary follows the protocols and remains undetected. Mostly attackers are semi-honest in nature.

### *B. Network view*

On the basis of the view of the network, an adversary is basically of two types: local and global. Local adversary can only view a part of the network at a glance whereas the global has the capability to view the whole network at a glance.

### *C. Resources*

An adversary has unlimited resources, computational power and memory. It has no dearth of resources. It can save all the captured messages which can be used to identify the routing path. Adversary is passive i.e it can only overhear the message, it cannot harm the sensor nodes like destroying sensor nodes, compromising some nodes etc.

### *D. Information Exposed*

The information exposed to the adversary depends on the author perspective. Some adversary has only knowledge about the sink location while other may know about the methods used at the nodes. Initially, adversary will be found near the base station. From base station it will start its strategy to capture the source. Adversary is mobile i.e it can move from one position to another. It can move towards the immediate sender of the captured message.

## IV. SOLUTIONS TO PRESERVE SOURCE LOCATION PRIVACY

This section provides a summarized view of the strategies used for preserving the location of the source. Basically, the solutions are categorized into several strategies such as random walk, dummy data sources, cyclic entrapment,

geographic routing and so on. In random walk, a packet follows a routing path of randomly selected intermediate neighbors in the network. It looks similar to a pattern followed in rumor routing. In dummy data sources, some nodes generate dummy packets and transmit them within the network. These nodes transmit the packets at the same rate as real events which confuse the adversary, preventing them from distinguishing between real and fake events. In cyclic entrapment strategy, the concept of multiple fake data sources is used. These nodes form a loop and exchange messages within the loop as soon as any member starts the transmission. The main idea behind this concept is that if the loop is large enough, then the adversary should take a while to decide that whether it is moving in a cycle or in the correct direction while hop by hop tracing of the traffic. Similarly, rests of the strategies are also used for preserving the location privacy of the source.

Based on the different strategies we are providing different techniques to preserve the location of Source Node(s) in WSN. We discuss regarding four techniques in this paper which is defined and explained below those solutions are:

- **Random Walk**
- **Multiple Phantom Nodes**
- **Baseline Flooding**
- **STaR Routing**

## 1. Random Walk

The packet tracing path is made completely random using this approach for an adversary which performs hop by hop tracing and traffic analysis attacks. For further transmission, a node selects their neighbours based on their forwarding probability. As in pure random walk, each node has equal probability to be selected as intermediate node. Therefore, the neighbours also including them who have already forwarded the message can be selected as intermediate nodes for forwarding the message. Then, there may be a possibility that the message may loop in a cycle near the source node. For example, as shown in **Fig. 3** a message from path *a* can follow a random walk of *a-b-c-d-e* to deliver it to the node at the end of path *e*, which creates a cycle of *b-c-d*. If this cycle comes under the hearing range of an adversary then it would directly move to path *a* from path *e*. So, pure random walk was not very efficient and also showed increased energy consumption by nodes.

At first under this approach comes the phantom routing scheme (PRS). It consists of two phases: the random walk phase and the flooding phase. As pure random walk was not a proper way of finding intermediate neighbour, so in order to avoid the repetition of paths, it introduced an extended version of random walk termed as directed random walk. In this, a node divides its neighbours into two groups that are opposite in direction. For instance, one group has neighbours in North-East direction and another has neighbours in South-West direction. When a node senses any subject in its range, it becomes the source node. For forwarding the message to the sink about the presence of subject in its range, it selects a group based on a flip of coin and forwards the message to the randomly selected neighbour of the selected group. The node which receives the message selects a neighbour from the opposite group and forwards the message to it. So, this message is forwarded in a zigzag manner for *h* hops. After *h* hops, the random walk terminates and the last node that receives the message is termed as phantom source. The message is then flooded in the network by the phantom source towards the sink.

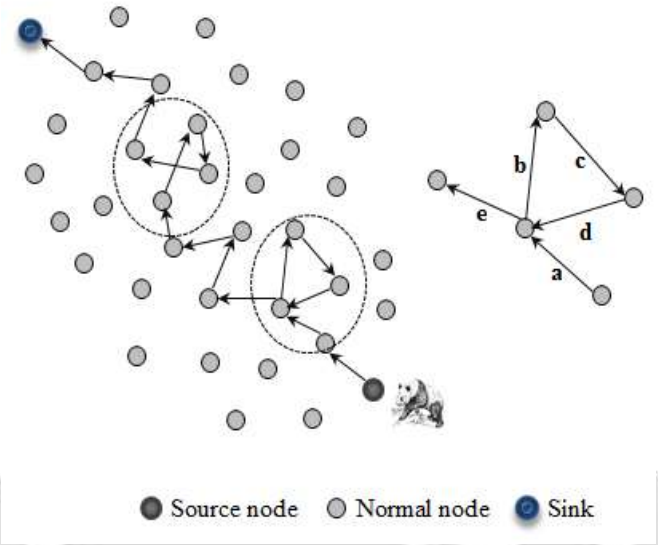



**Fig 3: Cycle in Pure Random Walk**

## 2. Multiple Phantom Nodes

In this technique we consider the network model as homogeneous sensor nodes deployed randomly that can behave as source node, intermediate node or phantom node. All these sensor nodes are static in nature which means they cannot move in the network.

**Source Nodes** : Sensor node that senses any event and forwards the message to the base station is called as source node.

**Intermediate Nodes** : Nodes that forward the received message towards its destination. When a sensor node senses any event then it forwards the message towards the sink with the help of neighbor nodes. These all neighbor nodes are called intermediate nodes.

**Phantom Node :** Sensor node that forwards the message of source node with its own identity.

All three nodes are similar in nature but performing different task at different time.

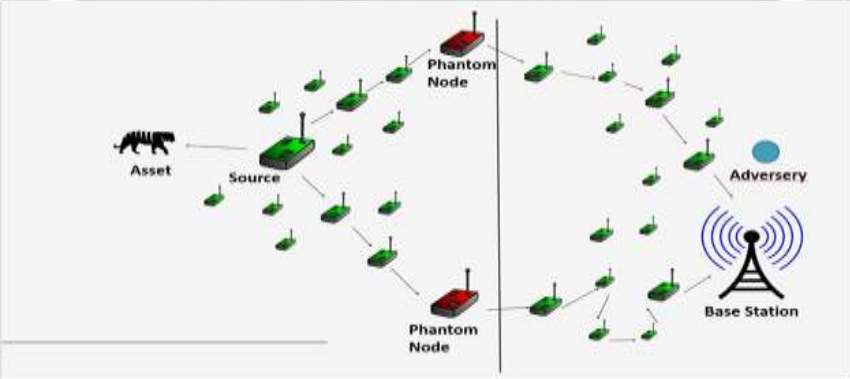



**Fig 4: Multi Phantom Routing Architecture**

Our proposed scheme consists of two phases:

(i) configuration phase (involves neighbor discovery, flooding node reports its hop count from BS and triplet section.

(ii) working phase (involves random walk and phantom selection based on given criteria).

- **Configuration Phase**

  During configuration phase, initially sink starts flooding with a message setting counter zero. Each node stores the counter value with sender I D. After that it forwards the message to its neighbor with

incremented value of counter by one.In this way,each sensor node has well knowledge that base station is how much hop distance away. After that each node informs the hop-distance to the sink. Sink maintains a hop distance table from where it creates set of triplets of sensor nodes. In a triplet, each sensor node behaves as phantom node for other two nodes. A triplet is selected in such a way that no two sensor nodes and sink are co-linear and the angle between each two node with the sink should be atleast 30 degree. If sensor nodes are co-linear with the sink then the source node would lie in the path between phantom node and sink. When this condition arises, then the privacy can be easily breached by the adversary. During configuration phase, it is assumed that all sensor nodes have been localized and the sink node has well knowledge of all the sensor nodes. Now base station randomly chooses three sensor nodes that are nearly same hop distance away and then calculates the angle between them.

- **Working Phase**

  After the completion of configuration phase, the working phase starts during which the communication between the source node and the sink node is performed. Triplet selection has been done during configuration phase. Each node has ID of two different sensor nodes that are in triplet. These two sensor nodes will behave as phantom node one at a time. When a source node senses any event then it randomly generates a number within 1 to 10. If the generated number is greater than 5then the first node is selected otherwise second node is selected as phantom. After selecting the phantom node, source node forwards the message to randomly selected neighbor with phantom node as destination. Selected neighbor forwards the message towards the destination phantom node with shortest path algorithm. It also includes its ID in message content. After receiving the message at phantom node, it checks the sender of the message. If source is its phantom node then it forwards the message towards the sink with its own ID by using shortest path algorithm. This we can explain with the help of fig A. Here, source first randomly generates a number a between 1 and 10. After that it is checked whether a is greater than 5 or less than equal to 5. If the a is less than or equal to 5 then the PI is selected as phantom node otherwise P2 is selected as phantom node. After the selection of phantom node, source randomly chooses x from set of neighbors N. Now, source passes message M to x with P as destination. Then, after some intermediate nodes message reaches to the phantom P with the help of shortest path algorithm. Now phantom node checks whether source is its phantom. If the condition satisfies then the P forwards the message to the sink with the help of shortest path algorithm. This phase is shown in figure below.

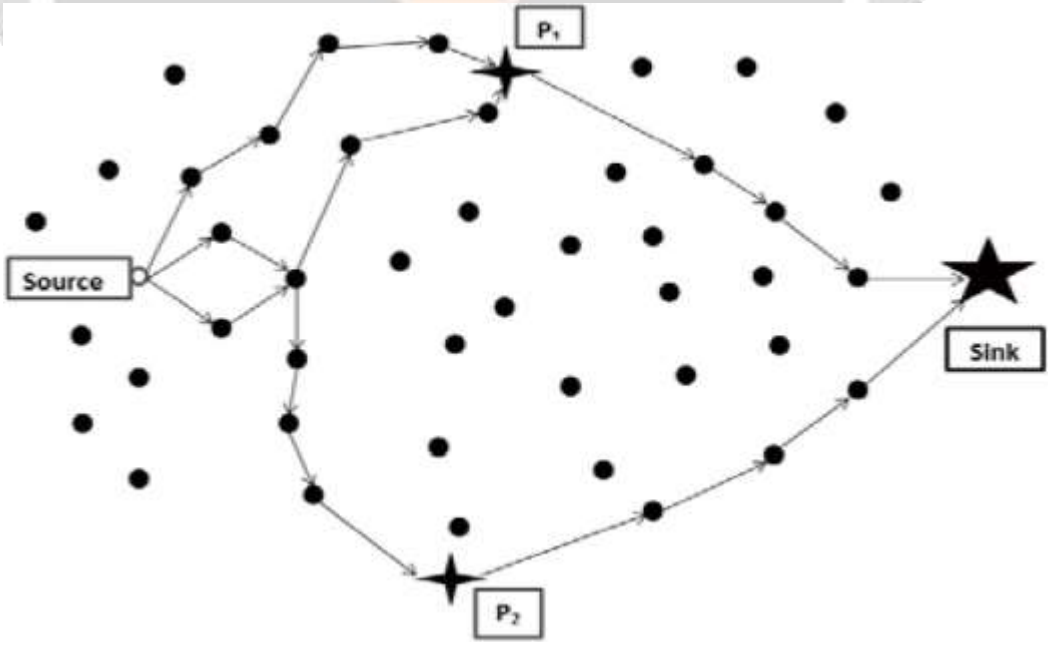



**Fig 5: Working Phase in Multiple phantom scheme**

**3. Baseline Flooding**

In baseline flooding, source node sends the packet of information to its neighbors and neighbors re-transmit the packet to its immediate neighbors so there will be many paths created for a packet to reach the to destination from the source so it is very difficult for the adversary to track the location of the source node. Once the packet is sent from source to its neighbors, these neighboring nodes will not retransmit the packet back to the source node. This is where one can provide the privacy for a location of the node. The **Fig 6** below shows the actual working of this technique.

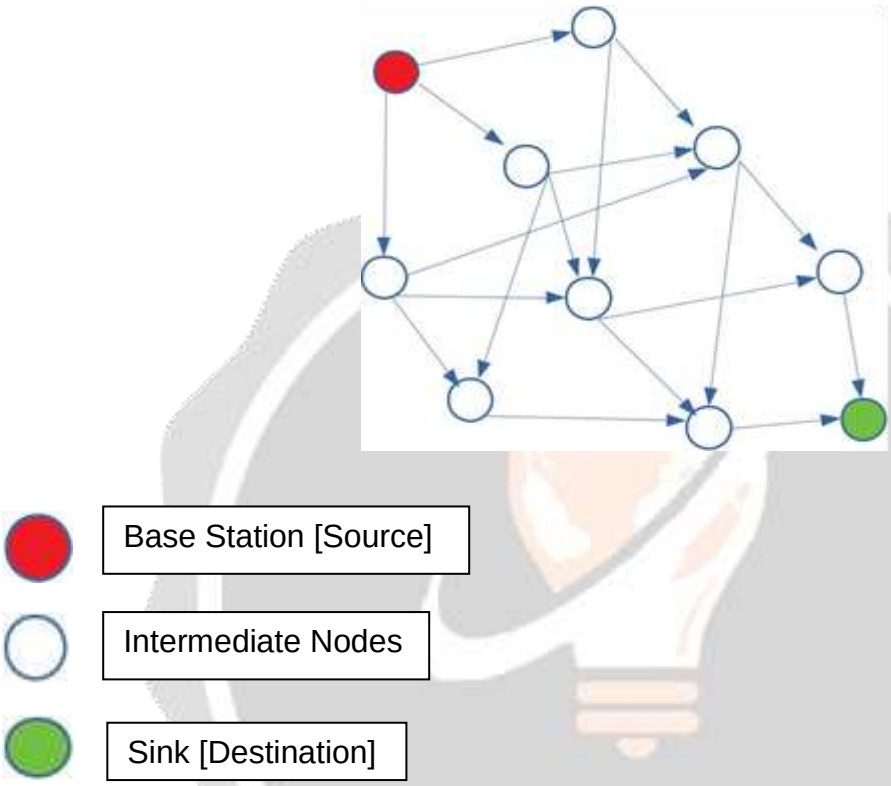



**Fig 6: Baseline Flooding**

The efficiency of a baseline flooding mechanism is based on the number of messages the source has sent before this source node is located by the adversary. This baseline flooding mechanism provides enough privacy to the source node. It can be only of the adversary is local one. In future we can implement global one. The power consumption of this mechanism is more due to the flooding of the data in the network. One most important advantage of this mechanism is that there is guaranteed data arrival to the base station. Delay also very less since there is no factors that affect the transmission speed of the packet in the network.

Design steps:

1. We are creating the source node, intermediate nodes and the destination node.
2. Creating the link between all neighbor nodes for the packet transmission from source to destination.
3. Then we send the UDP packet from source to destination according to the baseline flooding mechanism. Here we use the UDP packets because there is no retransmission occur in the UDP.

So, backtracking is very difficult from the sink to source but it is possible, it take more time for the attacker. Here we are design for the local adversary, i.e it contain limited number of nodes and it is the small part of the network.

**4. STaR Routing**

We introduce this technique as the Sink Toroidal Region (STaR) routing. With this technique, the source node randomly selects an intermediate node within a designed STaR area located around the SINK node. The STaR area is large enough to make it unpractical for an adversary to monitor the entire region. Furthermore, this routing protocol ensures that the intermediate node is neither too close, nor too far from the SINK node in relations to the entire network. While ensuring source location privacy, our simulation results show that the proposed scheme is very efficient and can be used for practical applications. The main idea is to, first, route the message to a node away from the actual message source randomly, then forward the message to the SINK node using single path routing. However, both theoretical and practical results demonstrate that if the message is routed randomly for *h* hops, then the message will be largely within *h*/5 hops away from the actual source. In this scheme, the source node first randomly selects an intermediate node at the sensor domain based on the relative location of the sensor nodes. The intermediate node is determined by two factors: 1) it must be outside the constrained region around the source and 2) it is normally distributed outside the constrained area. With this method, the selected intermediate node is expected to be away from the real source node, which provides local location privacy. In order to provide both local and global location privacy over the sensor network, the selection of intermediate nodes has to be totally random, i.e., every sensor node in the network has the same probability of being selected as the intermediate node for any source node. Unfortunately, the energy consumption for this design is quite high. In this paper, a design tradeoff has been made to balance these two needs. The intermediate nodes are evenly distributed in the STaR so that the messages can be routed to the SINK node from all possible directions. The distribution of STaR area is given in below figure.

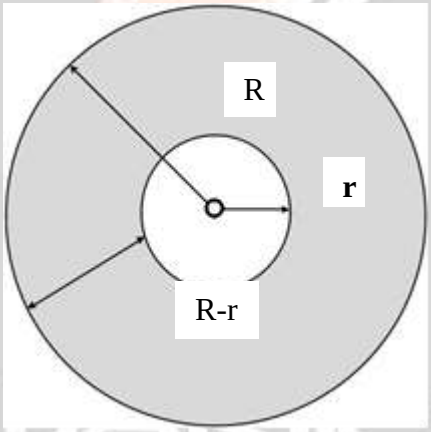



**Fig 7: Distribution of STaR Area**

We propose a two-phase routing protocol to provide source-location privacy. In the first phase, the source node randomly selects an intermediate node at the sensor domain and routes the message to the random intermediate node. The random intermediate node services as a fake source when the message is forwarded to the SINK node. In this scheme, the random intermediate node would be located in a pre-determine region around the SINK node. We call this region the Sink Toroidal Region (STaR). In the second phase, the intermediate node then forwards the message to the SINK node by single-path routing. The goal of the proposed scheme is to provide local and global source-location privacy with adequate energy-efficient routing. Local privacy is obtained by the fact that the intermediate node is expected to be neither too close nor too far away from the real source, for most cases. The STaR area would be a large area with at least a minimum radius distance *r* , from the SINK node to provide global privacy. Also, the STaR area guarantees that the intermediate node is at most a maximum distance *R* away from the SINK node to limit the energy consumption in the routing paths. This routing scheme is designed to give the illusion that the source node is sending messages to the SINK node from all the possible directions.

The following assumptions are made about the system:

• The network is divided into grids. The sensor nodes in each grid are fully connected. In each grid, there is one header node responsible for communicating with other nearby header nodes. The whole network is fully connected through multi-hop communications

• The SINK node is the destination location that data messages will be routed to. The information of the SINK node is made public. On detecting an event, a sensor node will generate and send messages to the SINK node through a multi-hop routing.
• Each message will include a unique dynamic ID where the event was generated. Only the SINK node can determine the source node location based on the dynamic ID.
• The sensor nodes are assumed to know their relative locations and the SINK node location. We also assume that each sensor node has the knowledge of its adjacent neighboring nodes. The information about the relative location of the sensor domain may also be broadcasted through this network for routing information update

## V. CONCLUSION

Source location privacy is a serious issue for many monitoring and remote sensing applications. Source-location privacy is vital to the successful deployment of wireless sensor networks. In many scenarios, an adversary may be able to trace back to the source location if not handled properly. In this paper, we have proposed four techniques for source location privacy. The Multiphantom routing protocol to confuse the adversary by creating alternate paths from source to sink. This protocol also keep in mind the energy issues of WSN and avoids the use of dummy packets and flooding in working phase. The proposed protocol works better than single phantom based approach. Future work may be done on analyzing the performance with respect to increase of phantom nodes. we have concluded that most of the solutions based on Random walk will only prove to be effective against the local adversary. This applies that they are not effective against global and multi-local adversary. We first discussed the background view of the source location privacy and the adversarial capabilities. Then we explained the solutions that we have found for preservation of a source node through baseline flooding, this mechanism developed based on the factors like efficiency, delay, power consumption etc. Source-location privacy is vital to the successful deployment of wireless sensor networks. In this paper, we then introduced a STaR routing scheme for local and global source-location privacy protection. Our simulation results demonstrate that the proposed STaR routing scheme can achieve excellent performance in energy consumption and delivery latency. Message delivery ratio is slightly lower than the other schemes but it is still satisfying overall.

## REFERENCES

[1] Prabhat Kumar, J.P Singh, Prateek Vishnoi and M.P Singh, "Source Location Privacy using Multiple-Phantom Nodes in WSN", TENCON 2015 - 2015 IEEE Region 10 Conference, Macao, China, 10.1109/TENCON.2015.7372969.
[2] Shruti Gupta, Bhaskar Prince "Preserving privacy of source location using random walk: A survey", 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 10.1109/RTEICT.2016.7808199.
[3] Leron Lightfoot, Yun Li, Jian Ren "Preserving Source-Location Privacy in Wireless Sensor Network using STaR Routing", Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE, Miami, FL, USA, 10.1109/GLOCOM.2010.5683603.
[4] Revati A. Parate, Pragati Patil, Girish Agarwal and Abha Gaikwad Patil "Survey On Location Privacy Preserving Schemes In Wireless Sensor Network". International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Vol. 1 Issue 9, November-2012.
[5] Shaik Rasool Saheb, CH. Raja Jacob "Security Preserving Protocols For Location Monitoring System in Wireless Networks". Rasool- International Journal of Computer Science information and Engg., Technologies ISSN 2277-4408 || 01062012-018.
[6] A.C. Priya Ranjani N. Swarna Jyothi B. Sriramulu "Location Privacy Methods in Wireless Sensor Networks". International Journal of P2P Network Trends and Technology (IJPTT) -Volume3 Issue3- April 2013.
[7] Pradeep Kumar Roy, Rimjhim "An efficient privacy preserving protocol for source location privacy in wireless sensor networks", IEEE Xplore Digital Library, DOI: 10.1109/WiSPNET.2016.7566305
[8] Jun Huang, Meisong Sun, Shitong Zhu, Yi Sun, Cong Xing, Qiang Duan "A source-location privacy protection strategy via pseudo normal distribution-based phantom routing in WSNs", ACM Digital Library, DOI: 10.1145/2695664.2695843.
[9] Chen Gu, Matthew Bradbury, Arshad Jhumka "Phantom walkabouts in wireless sensor networks", ACM Digital Library, DOI: 10.1145/3019612.3019732.
[10] Abhilash N, "Providing Source Location Privacy in WSN using Random Route Adoption Method", International Journal of Advanced Research in Computer Science & Technology.
[11] Rui S, Mayank Goswami, Jie Gao, Xianfeng Guhi, "Is random walk truly memoryless — Traffic analysis and source location privacy under random walks", IEEE Xplore Digital Library, DOI: 10.1109/INFCOM.2013.6567114.

.