JCSE International Journal of Computer Sciences and Engineering Open Access

Research Paper Vol.-7, Issue-2, Feb 2019 E-ISSN: 2347-2693

An Anonymity Region Construction and Two Fold Location Protection Scheme for Improving Source and Sink Location Privacy in WSN

R. Jayanthi[1*], M. Mohanraj[2]

[1]Computer Science Teacher, Sanjose Matriculation Higher Secondary School, Kattoor, Mettupalayam, Tamilnadu, India

[2] Department, Department of Computer Application, Dr SNS Rajalakshmi College of Arts and Science, Chinnavedampatti, Coimbatore, Tamilnadu, India

**Corresponding Author: jayanthiswamy2019@gmail.com*

DOI: https://doi.org/10.26438/ijcse/v7i2.724730 | Available online at: www.ijcseonline.org

Accepted: 12/Feb/2019, Published: 28/Feb/2019

Abstract—Generally, Wireless Sensor Networks (WSNs) refers to a spatially distributed autonomous network which consists of several sensor nodes. These nodes sense and transport the data to the sink node through adjacent nodes. Contextual privacy is the main challenging issue in WSN. The attacker node gathers information from traffic patterns between source and sink. The protection of source and sink location is very essential to prevent the attacker from gathering information. An all-direction random routing algorithm (ARR) was proposed to protect source-location from attacker node. This algorithm utilized agent nodes to establish a path between sources and sink nodes using only local decisions. ARR is efficiently protecting source location but it exposes direction information of sink. So in this paper, ARR is improved by injecting fake packets and random walk of real packets to hide direction information. Additionally, the anonymity of source and sink location is improved by Two-fold location privacy protection scheme where anonymity is constructed around the source and sink node based on geographic information to hide actual location. In anonymity region, packets are sent from a fake source node and received by fake sink node. The number of fake source and sink is selected based on traffic flow.

Keywords— Wireless Sensor Network, Contextual privacy, All-direction Random Routing algorithm, Improved All-direction Random Routing algorithm, Two-fold location privacy protecting.

I. INTRODUCTION

Wireless Sensor Network (WSN) [1,2] are spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, pressure, etc. and to cooperatively pass their data through the network to the main location. WSN [3] composed of two kinds of nodes is the sensor node and the sink node. The sensor node and the sink node behave as an information source and an interface respectively [4]. The sink node is prone to various kinds of attacks due to the WSN's nature. Various kinds of attacks may compromise different privacy problems in WSNs. The privacy is classified into categories as content privacy and contextual privacy.

Content privacy problem [5] means that adversaries can manipulate and observe the content of packets. Such problems are effectively handled by encryption and authentication mechanism. Even after strong encryption and authentication mechanisms contextual information about traffic in the network still exposed. For contextual privacy, adversaries can get sensitive information by analyzing traffic patterns and eavesdropping the network traffic instead of monitoring the content of packets. Protecting the source locations of the valuable packages is considered as the most important challenges in WSN.

An all-direction random routing algorithm (ARR) [6] was designed to protect the source-location. The routing process of ARR was composed three stages are choosing an efficient agent node, delivering the package from the source node to the agent node and took it from the agent node to the final destination. Instead of forming a path between the source and sink with knowing the entire topology of the network, this approach is used agent nodes for establishing a path between the source and sink nodes and using only local decisions. This approach is efficiently protecting the source location. However, this approach exposed direction information of sink.

So in this paper, an Improved ARR is proposed to improve ARR which hides direction information by injecting fake packets and random walk of real packets. In addition, the anonymity of source and sink location is enhanced using two-fold location privacy protection scheme. In this approach, anonymity is constructed around the source and sink node by using geographic information for hiding actual location. Packets are sent from fake source node and received by a fake sink node in anonymity region. By using traffic flow, the number of fake source and sink is selected.

The rest of the article is structured as follows: Section II presents the literature survey related to privacy preserving in WSN. Section III explains the proposed methodology. Section IV illustrates the experimental results of the proposed protocol. Finally, Section V concludes the research work.

II. LITERATURE SURVEY

The problem of preserving the location privacy of sensors of a wireless sensor network [7] was addressed at the time the sensors sent a reply to a query broadcast by the Base Station (BS). It was dealt with one of the worst scenarios for privacy i.e. when sensors were queried by a BS to provide the MAX of their stored readings. The MAX was computed by a probabilistic and scalable protocol which had the following features are guaranteed the location privacy of the sensors replying to the query, resilient to an active adversary to alter the readings sent by the sensors and allowed to trade-off the accuracy of the result with the overhead increase. Sometimes, the solution for the problem of preserving the location privacy is not viable.

In order to guarantee the privacy of the node's location and the event, a differentially private branching framework [8] was presented. It was based on the principle that an event was generally monitored by multiple nodes which leads to low sensitivity to transmission. Additionally, fake traffic was required to be generated when an event was reported by a small number of nodes. The backtracking was prevented by using dummy sources. The privacy of an event also inflicted the constraint that an adversary must not be able to distinguish the fake and real traffic. However, this framework has a high computational complexity.

The source location privacy (SLP) [9] was provided by presented Dynamic Single Path Routing (DynamicSPR) algorithm. The intractable nature of SLP was addressed by the static heuristic in DynamicSPR. It was a hybrid approach which presented to circumvent this issue. It utilized a directed random walk for allocating fake sources. However, the delivery ratio decreased due to the higher number of messages and the increased likelihood of occurrence of the hidden terminal problem.

A realistic semi-global eavesdropping attack model [10] was proposed to preserve the source location in WSN. This model measured source location privacy by describing α-angle anonymity against the semi-global eavesdropper. In order to preserve α- angle anonymity, a Mule-Saving-Source (MSS) was designed by adapting the function of data mules. The total delay was reduced by reducing the buffering time at mule and source.

A distributed solution for source location privacy [11] was proposed using the Fake Source and Phantom Routing (FSAPR) protocol. In this protocol, every time the source node sent a packet it was encrypted with a key which was already shared by the BS. Furthermore, each message contained the identity along with the sensed data which was encrypted with a key which was shared with the BS. It protected the location information of a sensor node sensing an event and sending it to the BS.

III. METHODOLOGY

In this section, the proposed Improved ARR (IARR) is described in detail where ARR is improved by injecting fake packets and random walk of real packets for hiding direction information. Moreover, the anonymity of source and sink location is enhanced using two-fold location privacy protection scheme. In this approach, anonymity is constructed around the source and sink node by using geographic information for hiding actual location. Packets are sent from a fake source node and received by fake sink node in anonymity region. By using traffic flow, the number of fake source and sink is selected. The flow of the proposed IARR approach is depicted in Figure 3.1.

Initially, node A and B chose two numbers in the range $[0, m - 1]$ randomly and insert the fake packets, i.e.,$r_A \; and \; r_B$ respectively, and they calculate $T_A = \alpha^{r_A} \; and \; T_B = \alpha^{r_B}$ based on $r_A \; and \; r_B$. If node A wants to verify the legal identity of node B, node A needs to request the certificate from node B and then it can verify the legality of node B by checking whether $P_{TA}(ID_B \| P_B, sig_B) = true$. Similarly, node B can check the legality of node A. If at least one node of A and B is a parasitic node, the process ends. The measure of anonymity is calculated based on geographic location. The node A and B generate the shared key K_{AB} and K_{BA} when both A and B are legal nodes of the network. Node A can compute $K_{AB} = h(P_B^{r_A} || T_B^{S_A})$ and node B can compute $K_{BA} = h(P_A^{r_B} || T_A^{S_B})$. Considering that $P_A = \alpha^{S_A}, T_A = \alpha^{r_A}, P_B = \alpha^{S_B} \; and \; T_B = \alpha^{r_B}$, it is ascertain that $K_{AB} = h(\alpha^{S_B r_A} || \alpha^{r_B S_A}) = K_{BA}$ and the two nodes are able to successfully construct the shared key. In addition, hide the direction information and packets are sent from a fake source node and received by the fake sink node in anonymity region. By using traffic flow, the number of fake source and sink is selected.

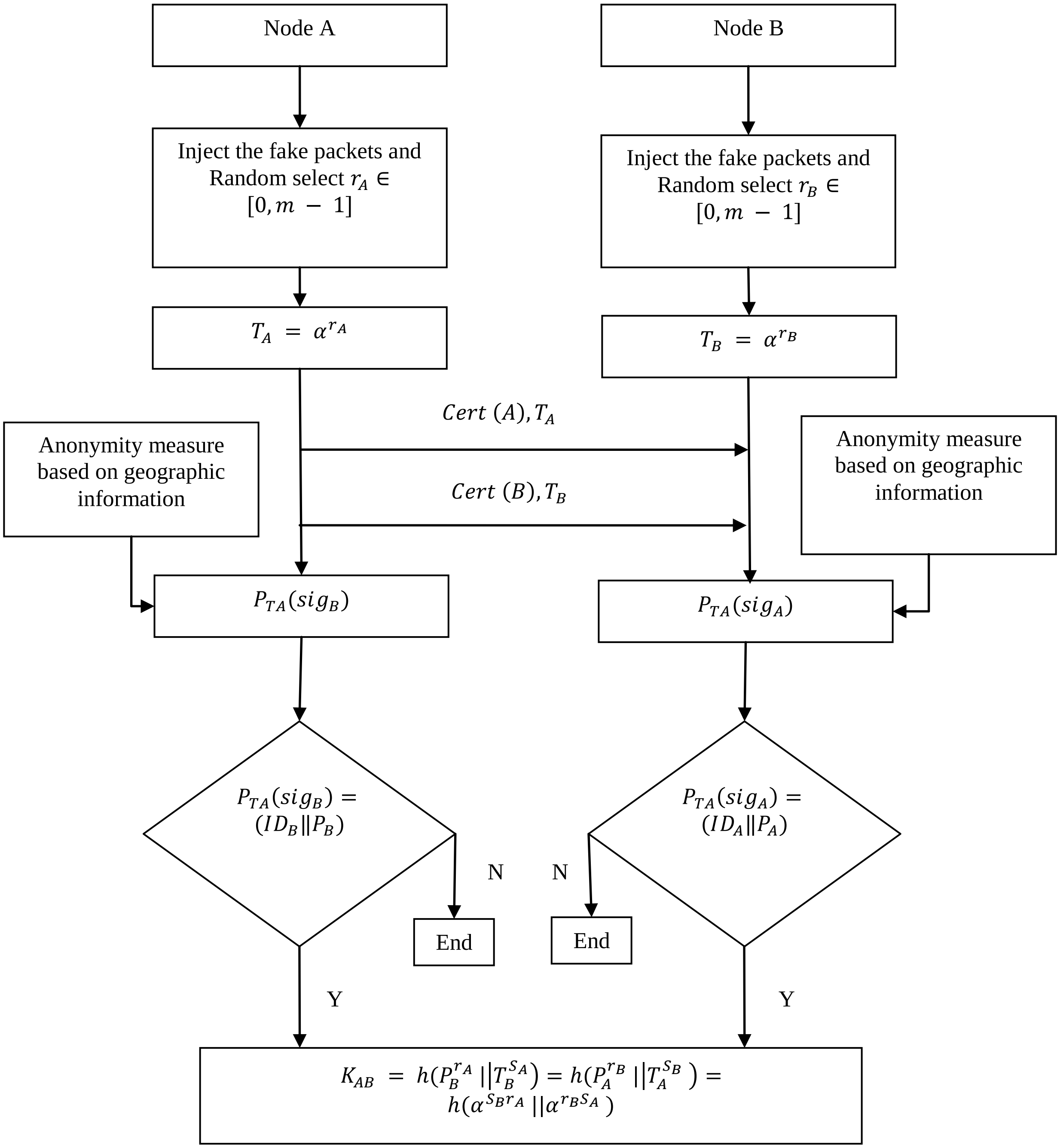
Node A
Node B
Inject the fake packets and Random select $r_A \in [0, m - 1]$
Inject the fake packets and Random select $r_B \in [0, m - 1]$
$T_A = \alpha^{r_A}$
$T_B = \alpha^{r_B}$
$Cert\ (A), T_A$
$Cert\ (B), T_B$
Anonymity measure based on geographic information
Anonymity measure based on geographic information
$P_{TA}(sig_B)$
$P_{TA}(sig_A)$
$P_{TA}(sig_B) = (ID_B \| P_B)$
$P_{TA}(sig_A) = (ID_A \| P_A)$
N
N
End
End
Y
Y
$K_{AB} = h(P_B^{r_A} \| T_B^{S_A}) = h(P_A^{r_B} \| T_A^{S_B}) = h(\alpha^{S_B r_A} \| \alpha^{r_B S_A})$

Figure 1. Flow of the Proposed Approach

A. Network Model

The network is divided into squares with the same size and then the sink nodes are deployed manually on the vertices of the squares. These sink nodes are logically equivalent and the source node can send the packages to any one of the sink nodes. Then, by using either wired or wireless channels each pair of sink node can communicate with each other directly and they can share information. Different types of targets in the network are determined by using a k-nearest neighbor. When no target is detected, every node processes a sleeping schedule and keeps silent. However, if a node detects a target in its duty regions, it needs to remain active until the target moves out of its duty regions. Once a target is detected by some nodes, these corresponding nodes immediately and accurately locate the target in a cooperative manner and send the information of the target to any one of the sink nodes.

Based on the contextual information of the WSNs, the adversaries try to locate the source nodes of the packages. The information of traffic distribution is obtained by the adversaries deploy various complicated parasitic nodes with supporting equipment such as communication models and spectrum analyzer. Initially, the parasitic nodes are uniformly deployed around the sink nodes. Assume that the locations of targets are uniformly distributed in the whole network and all the destinations of the packages are the sink nodes. If a parasitic node observes that a package is sent from node n_i, it moves to node n_i and waits until it hears another package being sent from node n_j. Then, the parasitic node moves to node n_j and repeats the process until it reaches the place near to the source node. Further, we assume that the parasitic nodes can communicate with each other and make decisions in a collaborative way.

B. Improved ARR

The ARR is improved by injecting fake packets and random walk of real packets for hiding direction information. It is performed three phases,

- The initial phase is the period before a real packet reaches its first intersection node, in this phase, this real packet is routed along the shortest path.word “data” is plural, not singular.
- The second phase is the period between a real packet reaches the first intersection node and finishes the random walk. The first intersection node should send N_fake (the number of fake sinks,) fake packets to N_fake fake sinks for the real packet. This real packet is transmitted continually with the probability of ph, and when a node decides to transmit this real packet, it selects the next hop from the further neighbor list with the probability of pf, and from the closest neighbor list with the probability of 1-pf. What’s more, when a node transmits a real packet, it should inject N_fake fake packets and forward these packets to random destinations, which is used to hide the real packet path.
- The third phase is the period between a packet finishes the random walk and reaches the sink. In this phase, the real packet is routed along the shortest path like the first phase. When this packet reaches its ith (i>1) intersection node, the intersection node should send N_fake fake packets to random destinations for it.

C. IARR with anonymity region (IARR-A)

In the proposed approach, the anonymity of source and sink location is enhanced by using a two-fold location privacy protection scheme. In this approach, anonymity is constructed around the source and sink node by geographic information to hide the actual location. In anonymity region, packets are sent from a fake source node and received by fake sink node. Once a sensor i reports its measurement to the sink, it encrypts the message with its symmetric key K_i and forwards the packet along a random path. Unlike many existing routing algorithms, the location or ID of the sink is not included in the packet. The advantage of this approach is to avoid the attackers from obtaining the destination of the packet even they can capture the intermediate nodes and read the packet.

Since i does not know the location of the sink, it forwards the packet randomly to any of its neighbors. When the next hop j receives the packet, it again forwards the packet to one of its neighbors k randomly and increases the hop count field H in the packet by one. The hop count field H in the header of the packet is initialized to zero by the source node. It indicates the number of hops that the packet has traveled. The above forwarding process repeats hop-by-hop until H = L, where L is the pre-defined length of the random path. Note that the packet will continue traveling in the network even it has already reached any of the sinks. Similarly, it is possible that the packet has never visited any sink at the end of its travel.

More specifically, node i sends the packet in this format $< i|Y_{type}|H|Y_{K_i} >$, where Y_{type} is the type of message in the packet, Y_{K_i} is the message encrypted by symmetric key K_i of node i, and H is the number of hops traveled by the packet. The message type Y_{type} allows the sink to recognize the content of the packet. The sink will only decrypt the packet that contains messages of its interest. A packet may store the ID of the nodes that it has visited, such that the following intermediate nodes can avoid re-visiting them. This mechanism increases the chance for the packet to reach the sink as one can visit more different nodes. It can be achieved by concatenating the ID of the intermediate nodes to the packet, i.e. $< i\left|Y_{type}\right|H\left|Y_{K_i}\right|ID_1|ID_2| \ldots |ID_H >$, where $ID_1, \ldots, ID_H$ are the IDs of the nodes being visited. Moreover, instead of sending the packet along a single path, the packet can be delivered by multiple paths to increase its chance to reach the sink. For instance, the source node may send the packet to M neighbors, then these neighbors will forward the packet along different random paths independently.

Based on traffic flow, the number of fake source and sink is selected. To reduce the starkness of pronounced paths, the shortest path (SP) routing approach is modified by having each node selects one of the multiple parent nodes to route data to the base station. When a node needs to forward a packet, the node randomly selects one of its parent nodes to forward the packet. We call this scheme multiparent routing (MPR). Two methods are used for setting up multiple parents for each node. In the first method, the beacon message sent by the base station contains a level field. The base station sets the value of level to 0. When a node forwards a beacon message, it increments it by 1. So the value of level represents the number of hops that a node is from the base station along a particular path. A sensor node s selects all neighbor nodes whose level value is less than s's level value as its parent nodes. In the second method, a node monitors all beacon messages it receives before forwarding the first beacon message. Since a node s has to wait for some amount of time before forwarding a beacon message (waiting time in MAC layer), it selects all nodes from whom it receives a beacon message while waiting to forward the first received beacon message as its parent nodes.

An adversary has several ways to attack these multiparent routing setup schemes. A malicious node can claim a low-level value to attract other nodes and can use unfair media access control mechanisms to occupy the wireless channel. However, protecting routing schemes is beyond the scope of this paper. Here we assume that the routing set up scheme is relatively fast, so an adversary doesn't have enough time to attack routing set up the process. In addition, we use random walk technique, fractal propagation, Fractal propagation with different forking probabilities and Enforced fractal propagation for calculating the performance of traffic analysis [12].

IV. Results and Discussion

In this section, the efficiency of the ARR and IARR approach is tested in terms of average time delay, average energy consumption, average amount of data transmission, source detection probability, and sink detection probability. The simulation parameters are listed in Table 4.1.section provides an overview of the advantages and disadvantages of various decision tree techniques.

Table 1. Comparison of different decision tree techniques

Parameter	Value
Size of the network	400×400 m
Number of Nodes	10,000
Number of Sinks	16
Number of Targets	1
R_c	30 m
Adversaries hearing range	30 m
Number of parasitic nodes	N_p
V_1	$\left(\frac{d}{12}\right)^2$
Target monitoring scheme	k-nearest neighbors tracking
Event transmission rate	1 s
Length of data in package S	1024 bit
Length of the head of a package	32 bit

A. Average Time Delay

The source-sink distance in hops is defined as the number of hops when delivering a packet from the source node to the sink node through the proposed algorithm.

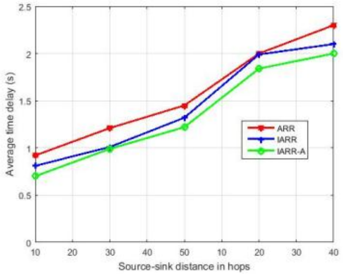

Figure 2. Comparison of Packet Delivery Ratio

Figure 2 shows the comparisons of proposed and existing techniques in terms of average time delay. The X-axis denotes Source-sink distance in hops and Y-axis indicates the average time delay value. When source-sink distance in hops is 40, the average time delay of proposed IARR-A is 13% less than ARR and 4.8% less than IARR. From this result, it is known that the proposed IARR-A has better average time delay than the IARR and ARR.

B. Average Energy Consumption

An Important concern in WSNs is energy consumption, which has a strong relationship with the amount of data transmission and the complexities of algorithms executed by the sensor nodes.

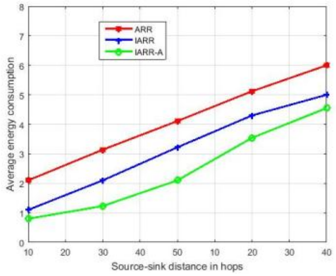

Figure 3. Comparison of Average Energy Consumption

Figure 3 shows the comparison of proposed and existing techniques in terms of average energy consumption. The X-axis denotes source-sink distance in hops. Y-axis indicates the average energy consumption value. When source-sink distance in hops is 50 the average energy consumption of proposed IARR-A is 24.2% less than ARR and 9% less than IARR. From this result, it is known that the proposed IARR-

A has better average energy consumption than the IARR and ARR.

C. Average Amount of Data Transmission

A round is defined as the whole process of monitoring a target, generating a packet and successfully delivering the packet to the sink node. All the data transmitted in the whole network are taken into consideration.

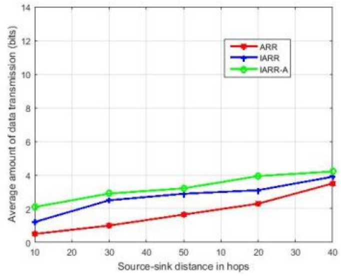

Figure 4. Comparison of Average amount of data transmission

Figure 4 shows the comparison of proposed and existing techniques in terms of the average amount of data transmission. The X-axis denotes source-sink distance in hops. Y-axis indicates the average amount of data transmission value. When source-sink distance in hops is 40 the average amount of data transmission of proposed IARR-A is 20.3% greater than ARR and 7.9% greater than IARR. From this result, it is known that the proposed IARR-A has a better average amount of data transmission than the other methods.

D. Source Detection Probability

The source detection probability is described as the probability that the parasitic nodes can locate the source nodes successfully.

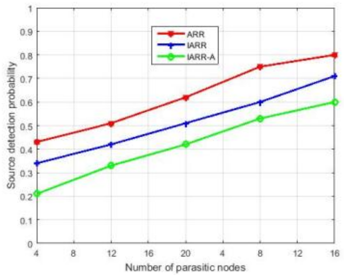

Figure 5. Comparison of Source detection probability

Figure 5 shows the comparison of proposed and existing techniques in terms of Lifetime. The X-axis denotes the number of parasitic nodes. Y-axis indicates the source detection probability value. When the number of the parasitic nodes is 16, the source detection probability of proposed IARR-A is 25% less than ARR and 15.5% less than IARR. From this result, it is known that the proposed IARR-A has better source detection probability than the other methods.

E. Sink Detection Probability

The sink detection probability is described as the probability that the parasitic nodes can locate the sink nodes successfully.

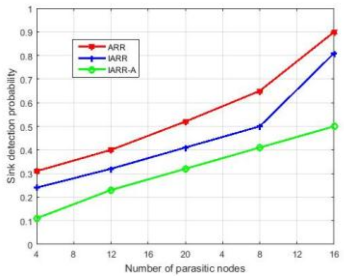

Figure 6. Comparison of Sink detection probability

Figure 6 shows the comparison of proposed and existing techniques in terms of Lifetime. The X-axis denotes the number of parasitic nodes. Y-axis indicates the sink detection probability value. When the number of the parasitic nodes is 16, the sink detection probability of proposed IARR-A is 44.4% less than ARR and 38.3% less than IARR. From this result, it is known that the proposed IARR-A has better sink detection probability than the other methods.

V. CONCLUSION

In this paper, direction information is hidden by using Improved ARR through injecting fake packets and random walk of real packets. By using two-fold location privacy protecting scheme, the anonymity of source and sink location is enhanced. In this scheme, the anonymity is constructed around the source and sink node using geographic information for hiding actual location. Packets are sent from a fake source node and received through fake sink node in the anonymity region. By using traffic flow, the number of fake source and sink is selected. The simulation results show that the proposed approaches are providing better results in terms of source detection probability, sink detection probability, average energy consumption, average time delay and the average amount of data transmission.

REFERENCES

[1] X. Yi, A. Bouguettaya, D. Georgakopoulos, A. Song, J. Willemson, “*Privacy protection for wireless medical sensor data*”, IEEE transactions on dependable and secure computing, Vol.**13**, Issue.**3**, pp.**369-380**, **2016**.

[2] K. Ravikumar, V. Manikandan, “*Detection of Node Capture Attack in Wireless Sensor Networks*”, International Journal of Scientific Research in Computer Science and Engineering, Vol.**6**, Issue.**4**, pp.**56-61**, **2018.**

[3] M. Poonam, Mahajan, "*WSN: Infrastructure and Applications*", International Journal of Scientific Research in Network Security and Communication, Vol.**6**, Issue.**1**, pp.**6-10**, **2018**.

[4] C. Gentili, G. Valenza, M. Nardelli, A. Lanatà, G. Bertschy, L. Weiner, P. Pietrini, "*Longitudinal monitoring of heartbeat dynamics predicts mood changes in bipolar patients: a pilot study*", Journal of affective disorders, Vol.**209**, pp.**30-38**, **2017**.

[5] L. Ranganath, K.S. Kavya, B.M. Priyanka, A.M. Shruthi, C. VidyaRaj, "*Security for Source Node Privacy in Wireless Sensor Networks*", International Research Journal of Engineering and Technology, Vol.**4,** Issue.**4**, pp.**1307-1309**, **2017**.

[6] N. Wang, J. Zeng, "*All-Direction Random Routing for Source-Location Privacy Protecting against Parasitic Sensor Networks*", Sensors, Vol.**17,** Issue.**3**, pp.**1-18**, **2017**.

[7] R. Di Pietro, A. Viejo, "*Location privacy and resilience in wireless sensor networks querying*", Computer Communications, Vol.**34**, Issue.**3**, pp.**515-523**, **2011**.

[8] B. Chakraborty, S. Verma, K.P. Singh, "*Staircase based differential privacy with branching mechanism for location privacy preservation in wireless sensor networks*", Computers & Security, Vol.**77**, pp.**36-48**, **2018**.

[9] M. Bradbury, A. Jhumka, M. Leeke, "*Hybrid online protocols for source location privacy in wireless sensor networks*", Journal of Parallel and Distributed Computing, Vol.**115**, pp.**67-81**, **2018**.

[10] M. Raj, N. Li, D. Liu, M. Wright, S.K. Das, "*Using data mules to preserve source location privacy in wireless sensor networks*", Pervasive and Mobile Computing, Vol.**11**, pp.**244-260**, **2014**.

[11] P.K. Roy, J.P. Singh, P. Kumar, M.P. Singh, "*Source location privacy using fake source and phantom routing (FSAPR) technique in wireless sensor networks*", Procedia Computer Science, Vol.**57**, pp.**936-941**, **2015**.

[12] J. Deng, R. Han, S. Mishra, "Countermeasures against traffic analysis attacks in wireless sensor networks", In 2005. SecureComm 2005. IEEE First International Conference on Security and Privacy for Emerging Areas in Communications Networks, pp.**113-126**, **2005**.

Authors Profile

Mrs Jayanthi.R pursued Master of Computer Science from Bharathiar University in 2009. She is currently pursuing M.Phil Dr SNS Rajalakshmi college of Arts And Science. And currently working as Computer Science Teacher in Sanjose Matriculation Higher Secondary School, Kattoor, Mettupalayam ,Coimbatore. Her Main Research work focuses on Advanced Networking. She has 6 years of Teaching Experience.

Dr Mohanraj M pursued **M.C.A., M.Phil., Ph.D.**, He is Research Guide and currently working as Assistant Professor in Dr SNS Rajalakshmi College of Arts and Science, Chinnavedampatti, Coimbatore. He had published more than 15 Journals. His Main Research work focuses on Networking. He has 12 years of Teaching Experience.