IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-6, June 2018 E-ISSN: 2347-2693

Provenance Based Mechanism to identify Packet Drop attack in WSN

Anudeepa[1*], Amutha S[2]

[1*]Dept.Computer Science, Dayananda Sagar College of Engineering, Bangalore, India
[2] Dept.Computer Science, Dayananda Sagar College of Engineering, Bangalore, India

e-mail: anuvkd94@gmail.com

Available online at: www.ijcseonline.org

Accepted: 18/Jun/2018, Published: 30/Jun/2018

***Abstract*—** Wireless sensor networks (WSN) is a wireless connection of the sensor nodes and they spontaneously build the networks so that the sensor data can be communicate wirelessly. Data collected from sensor network are used in decision making. A malicious adversary can introduce additional nodes in the network easily and tamper the information. Provenance is used for assessing the trustworthiness of the data, diagnosing network failures, detecting early signs of attacks, etc. But management of the provenance for sensor networks may lead to several challenging requirements, such as low energy and bandwidth consumption, efficient storage and secure transmission. We propose an efficient mechanism to securely transmit the data through the nodes using AODV Routing algorithm and secure provenance scheme to detect packet drop attacks in the network using Bloom Filter.

Keywords—Provenance, Ad-hoc On Demand Distance Vector(AODV)

I. INTRODUCTION

A sensor network refers to the group of small or tiny devices which includes the external battery power. Sensor networks are deployed in wireless infrastructure which monitors and records the environment conditions. The sensor network can connects to the Internet, wide area network (WAN) or local area network (LAN).The data collected from that is transmitted to the back-end systems for processing and analysis and they are used in WSN applications. Large number of sensor nodes will produce the huge data. The collected data are analyzed in the network at intermediate hops and base station (BS) have the capability of decision making. Data provenance is considered as the key factor for assessing data trustworthiness, since it summarizes the history of ownership and the actions performed on the data. BS trace the source and forwarding path of an individual data packet using data provenance. For each packet provenance must be recorded, but several important challenges arise due to the tight storage, energy and bandwidth constraints of sensor nodes. Therefore it is necessary to build a mechanism for provenance encoding and decoding with low over- head.

The proposed methodology depends on Bloom filters to encode provenance. The bloom filter includes two operations: test and add. Test is used to check whether a given element is in the set or not. Add is used to insert an element into the set. Bloom filter gives only false positive rate. If the bit is set to 1 then it is consider as its is in the set otherwise element is definitely not in the set. Existing system have two different transmission channels for provenance and data [1], we are using only a single channel for both data and provenance. Traditional provenance system use security analysis such as cryptography and digital signature[2] .We consider only message authentication code (MAC) schemes which is a fixed-size data structures.

II. EXISTING SYSTEM

Existing System is a novel lightweight scheme to securely transmit provenance for sensor data. This technique relies on in-packet Bloom filters to encode provenance. This research introduced efficient mechanisms for provenance verification and reconstruction at the base station. In addition, we extend the secure provenance scheme with functionality to detect packet drop attacks staged by malicious data forwarding node.

III. RELATED WORK

"Secure Network Provenance" describes why a routing table entry is present on a certain router, or where a given cache entry originated[9]. SNP allows the operators to track the track the faulty or misbehaving nodes using SNooPy. In SNP finding correct answers to forensic queries in an adversarial setting is difficult because the malicious nodes may lie to the querier. Forensic system requires trusted components but the components that are available today are not fully trusted ,so SNP helps to operate in a completely untrusted environment.

"The Case of the Fake Picasso: Preventing History Forgery with Secure Provenance" describes how to provide strong integrity and confidentiality assurances[2] for data provenance information and describe the provenance-aware system prototype that implements provenance tracking of data writes at the application layer, which makes it extremely easy to deploy. Uses intensively cryptography and digital signatures and they employ append-based data structures to store provenance, leading to prohibitive costs.

"Self-Identifying Sensor Data" introduce a technique to directly associate provenance information with sensor datasets [10]and it is similar to traditional water- marking but is intended for application to unstructured datasets. Algorithms are probabilistic in nature and are characterized by a combinatorial analysis. Unfortunately the realities of data usage complicate this, as data is passed around and used nth hand. The original source is often forgotten, without malicious intent.

"Secure Provenance Transmission for Streaming Data," a mechanism to securely transmit provenance for streaming data[11]. Provenance is extracted by the data receiver using an optimal threshold-based mechanism. This method minimizes the probability of provenance decoding errors. Experiments show this technique can recover provenance up to a certain level against perturbations to inter-packet timing characteristics.

"Towards Stateless Single-Packet IP Traceback" provides a lightweight single-packet IP traceback system which does not store any data in the network[12]. The proposed system relies on Generalized Bloom Filter which is a novel data structure called and also introduce the efficient improved path reconstruction procedure.

"Provenance-Based Trustworthiness Assessment in Sensor Networks" is a systematic method for assessing the trustworthiness of data items[4]. This approach uses the data provenance and their values for computing trust scores .Proposed a cyclic framework which well reflects the inter-dependency property. Key Distribution Mechanism is used to for provenance to asses the trustworthiness.

IV. PROPOSED METHODOLOGY

Proposed a provenance encoding method in which each node in the route securely transmit provenance information by using a Bloom filter (BF) that is broadcasted the each node in the path with the data. After receiving the data packets BS processed and conducts provenance verification and collection process. This technique also provide an extension to the provenance encoding scheme which supports the BS to find whether a packet drop attack has occurred in the path. Existing system have two different transmission channels for provenance and data, we are using only a single channel for both data and provenance. Traditional provenance system use security analysis such as cryptography and digital signature .We consider only message authentication code (MAC) schemes which is a fixed-size data structures.

A. Bloom Filter

Bloom Filter is a data structure which is used to test whether an element is a member of a set. Bloom filter requires less space and it is more efficient. It gives only the false positive results[3]. Bloom filter is an array of m bits (b0,b1,b2,……bm) that are initially set to 0. k independent hash functions (h1,h2….hk) used to describe how these bits are set or checked.

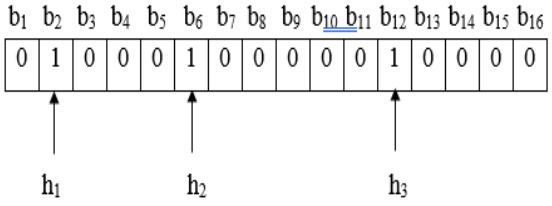

Figure 1. Bloom Filter with k=3

Once the insertion id done , the hash functions is applied and corresponding bits are set to 1. If any of the bit is set to 0 or output of the hash function is 0 then we can consider that the element is not in the set (no false negative property)[7]. The false positive probability can be obtained using the formula

$$fpb = (1-(1-1/m)kn)k \quad (1)$$

B.Provenance Encoding

provenance :Provenance (pd) is defined as a directed acyclic graph G(V,E)[5] where V is the vertex and E is the edges. For example consider a graph of leaf node nl , data packet d, v is the vertex.

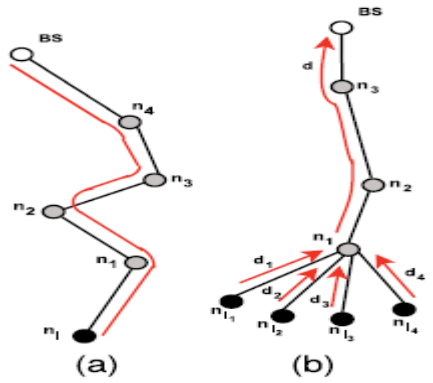

Figure 2. Provenance graph

For the figure (a) provenance generated by d is pd = < vl,v1,v2,v3,> . In figure (b) the node n1 generates the provenance aggregated data and they form the provenance graph as pd = < {vl1,vl2,vl3,vl4}, v1,v2,v3 >.

Provenance encoding is generating the vertices in the provenance graph, then insert it to the Bloom Filter using the equation .

$$Vidi=generate\ VID(ni,seq) = EKi\ (seq) \qquad (2)$$

The VID is generated per-packet based on the packet sequence number(seq) and the secret key Ki of the host node. E is the secure block cipher. Once the original message is transmitted to the network it has to be encoded, without encoding the message there is a chance of attacking the message by the attacker. So message should be encoded before sending into the network. Upon receiving the data packet intermediate nodes performs the data as well as provenance aggregation as shown in figure .Now each node computes bloom filter from children that is partial aggregated provenance[4].

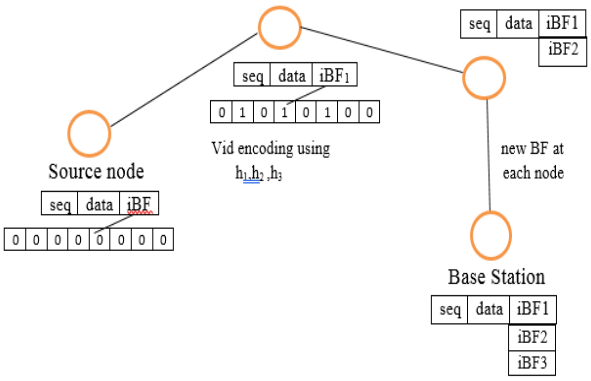

Figure 3. Provenance Encoding

Figure shows the provenance encoding in the each node. In the source node initially bloom filter is set to 0 and then hash function is applied to the elements in set . Source node will generates the Vid and inserts it to the Bloom filter. Then it is sent to the next node ,here again the Vid is generated and iBF1 is obtained .Hence in every node new BF is generated and transmitted to the Base Station.

C. Route Discovery

After the provenance encoding the path has to be found to send the encoded message to the base station. Routing table gives the forwarding information about the each node. But maintaining routing table at each node will lead to more usage of bandwidth and network resources. In our research we have used AODV protocol to overcome this problem.

AODV Routing :

AODV is an on-demand routing algorithm, that establishes paths only upon demand by source nodes. It maintains the established paths and uses only when they are need. Nodes that do not participate in active path, neither maintain any routing information nor participate in any periodic routing table exchange.

A. Provenance Decoding

The base station receives a packets from source which will be in encoded form. The first step is to verify the packets, once the verification process is clear then the base station knowns the data path and checks the in packet bloom filter to see the correct path has been followed. If base station doesn't know the data path then provenance collection process is performed, now the base station verifies its knowledge of provenance[8].

Algorithm 1 ProvenanceVerification

Input: Received packet with sequence *seq* and iBF *ibf*.
Set of hash functions H, Data path $P' = < n'_{l_1}, ..., n'_1, ..., n'_p >$

$BF_c \leftarrow 0$ // Initialize Bloom Filter
for each $n'_i \in P'$ **do**
  vid'_i = generateVID (n'_i, seq)
  insert vid'_i into BF_c using hash functions in H
endfor
if $(BF_c = ibf)$ **then**
  return true // Provenance is verified
endif
return false

Algorithm 2 ProvenanceCollection

Input: Received packet with sequence *seq* and iBF *ibf*.
Set of nodes (N) in the network, Set of hash functions H

1. Initialize

   Set of Possible Nodes $S \leftarrow \emptyset$
   Bloom Filter $BF_c \leftarrow 0$ // To represent S

2. Determine possible nodes in the path and build the representative BF

   for each node $n_i \in N$ **do**
   vid_i = generateVID (n_i, seq)
   if (vid_i is in ibf) **then**
   $S \leftarrow S \cup n_i$
   insert vid_i into BF_c using hash functions in H
   endif
   endfor

3. Verify BF_c with the received iBF

   if ($BF_c = ibf$) **then**
   return S // Provenance has been determined correctly
   else
   return NULL // Indicates an in-transit attack
   endif

E. Detection of packet drop

We consider that the links on existing path may produce natural packet loss and malicious nodes may add or remove the packets in that path[6]. For each packet provenance record consists of node ID , vertex ID and the sequence number. Upon receiving the data packets base station will perform provenance verification process. If the information does not match with the existing iBF then its is considered as packet has been lost or link failure has occurred.

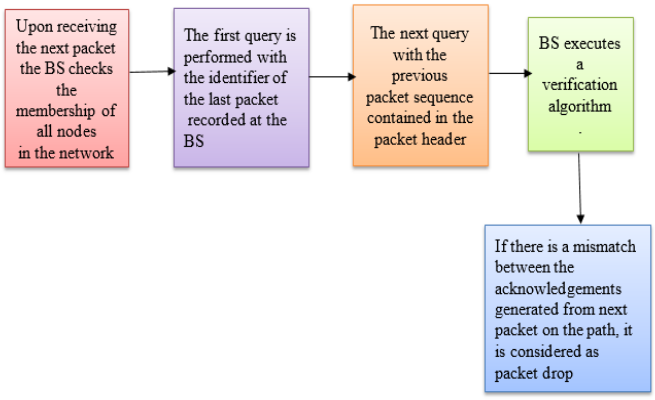

Figure 4. Detection of packet drop

V. PERFORMANCE ANALYSIS

1.We considered the generic secure provenance framework SProv [2] (R. Hasan, 2009) and lightweight version of this scheme is named as SSP, we have taken the provenance record at a node ni as Pi = < ni , hash(Di) ,Ci >, where hash(Di) is a cryptographic hash of the updated data, and Ci contains an integrity checksum as Sign(hash(ni, hash(Di)|Ci-1).

2. We also adapt a MAC-based provenance scheme, known as MP. Here a node transmits the nodeID and a MAC computed on it as the provenance record.

Detecting the packet drop :

Detection of the packet drop and identification of the malicious nodes can be done using provenance. the number of data packets transmitted by the source before reaching the converging condition is computed as follows[5].

$$N_{(\rho-1)} \frac{1}{(1-\alpha)^{\delta}} = \frac{\ln \frac{2}{\sigma}}{2\epsilon^2 \ (1-\rho-\epsilon)^{\delta}}$$

Where ρ is the natural packet loss value in hops.

VI. SIMULATION RESULTS

We implemented and tested the proposed techniques using the NS2 simulator. We consider a network of 39 nodes and vary the network diameter from 2 to 14. First, we compare our scheme with the SSP and MP in detecting provenance forgery and path changes. Next, we investigate the accuracy of the proposed method for detecting packet drop.

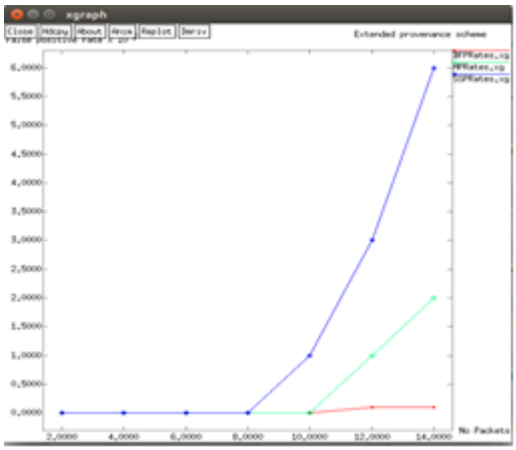

Figure 5. Proposed provenance scheme

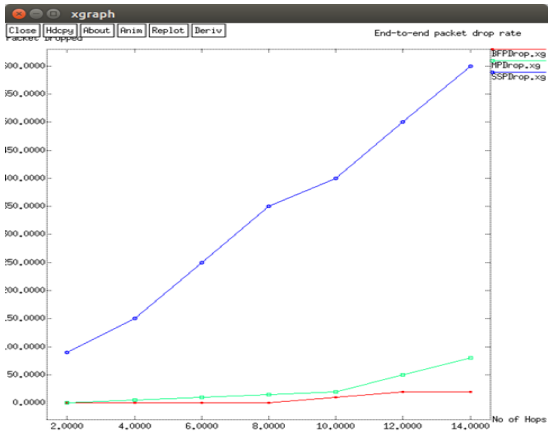

Figure 6. End to end packet drop rate

Figure 5 shows the performance of the extended provenance graph comparing with the SSP and MP model. Figure 6 describes the performance graph for end to end packet drop rate. Here generic secure provenance framework (SSP) has more packet drop in the network compared to the MAC based provenance scheme (MP). Here bloom filter has very less or no packet drops in the network.

CONCLUSION

Addressing the problem of securely transmitting provenance for sensor networks, and proposed a provenance encoding and decoding scheme based on Bloom filters. Bloom filters only encodes and decodes the data but it does not give a security for it. To secure the data in the network Asymmetric key cryptography is used, this scheme ensures confidentiality, integrity, non-repudiation, authentication of provenance. Extended the scheme to include packet sequence information that supports detection of packet drop in the network. Experimental results show that the proposed scheme is effective in security compared to generic secure provenance (SSP) and MAC-based provenance (MP).

REFERENCES

[1] Y. Simmhan, B. Plale, and D. Gannon, "A Survey of Data Provenance in E-Science," ACMSIGMODRecord, vol. 34, pp. 31-36, 2005.

[2] R. Hasan, R. Sion, and M. Winslett, "The Case of the Fake Picasso: Preventing History Forgery with Secure Provenance," Proc. Seventh Conf. File and Storage Technologies (FAST), pp. 1-14, 2009.

[3] C. Rothenberg, C.Macapuna,M.Magalhaes, F. Verdi, and A. Wiesmaier,"In-Packet Bloom Filters: Design and Networking Applications," Computer Networks, vol. 55, no. 6, pp. 1364-1378, 2011.

[4] H. Lim, Y. Moon, and E. Bertino, "Provenance-Based Trustworthiness Assessment in Sensor Networks," Proc. Seventh Int'l WorkshopData Management for Sensor Networks, pp. 2-7, 2010.

[5] Salmin Sultana, Gabriel Ghinita, Member, IEEE ,Elisa Bertino, Fellow, IEEE , and Mohamed Shehab,Member, IEEE Computer Society, "A Lightweight Secure Scheme for Detecting Provenance Forgery and Packet Drop Attacks in Wireless Sensor Networks", ieee transactions on dependable and secure computing, vol. 12, no. 3,may/june 2015.

[6] Salmin Sultana, "A Provenance based Mechanism to Identify Malicious Packet Dropping Adversaries in Sensor," in 31st International Conference on Distributed Computing Systems Workshops, 2011.

[7] A. Kirsch and M. Mitzenmacher, "Distance-Sensitive Bloom Filters," Proc. Workshop Algorithm Eng. and Experiments, pp. 41-50, 2006.

[8] S. Roy, M. Conti, S. Setia, and S. Jajodia, "Secure Data Aggregation in Wireless Sensor Networks," IEEE Trans. Information Forensics and Security, vol. 7, no. 3, pp. 1040-1052, June 2012.

[9] W. Zhou, Q. Fei, A. Narayan, A. Haeberlen, B. Loo, and M. Sherr, Secure Network Provenance, Proc. ACMSOSP, pp. 295-310, 2011.

[10] S. Chong, C. Skalka, and J.A. Vaughan, Self-Identifying Sensor Data, Proc. Ninth ACM/IEEE Intl Conf. Information Processing in Sensor Networks (IPSN), pp. 82-93, 2010.

[11] S. Sultana, M. Shehab, and E. Bertino, Secure Provenance Transmission for Streaming Data, IEEE Trans. Knowledge and Data Eng., vol. 25, no. 8, pp. 1890-1903, Aug. 2013.

[12] R. Laufer, P. Velloso, D. Cunha, I. Moraes, M. Bicudo, M. Moreira, and O.Duarte, Towards Stateless Single-Packet IP Traceback, Proc.32nd IEEE Conf. LocalComputer Networks (LCN), pp. 548-555,2007.

