Forgery and Packet Drop Detection Using BloomFilter Mechanism in Wireless Sensor Network

Shruthy H.N.
M.Tech, dept. of CSE
Bangalore Institute of Technology
Bangaluru, India
Email: shruthyhn@gmail.com

Nagamani D.R
Assistant Professor, dept. of CSE
Bangalore Institute of Technology
Bangaluru, India
Email: nmanirrr@gmail.com

***Abstract*— In Many application domains, large scale sensor networks are being deployed to collect sensor data that can be used in decision making for critical infrastructures. A malicious adversary may introduce a malicious node into the network or may compromise the existing legitimate node within the network. Hence, ensuring trustworthiness of data is necessary for effective decision making. Data provenance is a key factor in evaluating trustworthiness of data in sensor network. But, Provenance management in sensor network faces several challenges like low energy, storage and bandwidth consumption, limited resources, and adversary attack during transmission. In this paper, a novel lightweight scheme is proposed to securely transmit provenance data in wireless sensor network. The proposed technique uses in-packet bloom filter to encode provenance data. We introduce efficient mechanism for provenance verification and reconstruction of provenance at base station. Also the scheme is extended with additional functionality to detect packet drop attacks staged by consecutive malicious nodes, forwarding the data. We evaluate the proposed technique both analytically and empirically, and the results obtained using proposed scheme proves to be effective in detecting forgery and packet loss in multiple consecutive malicious sensor nodes.**

Keywords—wireless sensor network, provenance data, bloom filter, security.

I. INTRODUCTION

Sensor networks are becoming increasingly popular in numerous application domains such as environmental monitoring, medical, military surveillance, power grids, etc. Data produced at sensor source nodes are processed in network, at intermediate nodes on their way to base station to be analyzed for decision making. The diversity in data sources creates a need ensure the trustworthiness of the data so that, only reliable data is considered for decision making process. The data provenance is an effective method to evaluate data trustworthiness as it summarizes the history of owner ship and actions performed on the data. The provenance collection and transmission in sensor networks has not been addressed extensively. In this paper, the problem of secure and efficient data provenance transmission and processing is investigated for wireless sensor networks.

In multi-hop sensor networks, base station can use data provenance to trace the source and forwarding path to of an individual data packet in streaming data transmission. To achieve this, the provenance should be recorded for every packet. But several challenges arise because of small storage capacity, limited energy in sensor nodes and bandwidth consumption on sensor network [2]. Therefore, it is necessary to use a light-weight mechanism to obtain provenance data to reduce these overhead in the sensor network. Generally the sensor nodes deployed operates in an untrusted environment where they can be subjected to adversary attacks [5], [3]. Hence, it is necessary to address security requirements such as confidentiality and integrity of the provenance data. The aim is to design a provenance encoding and decoding mechanism, that fulfils the security and performance requirements. S Sultana and G Ghinita [1] have given a scheme to binding data and provenance together but, it limits to only single malicious node.

The proposed provenance encoding strategy securely embeds provenance information of each node, on the path traversed by the data packet, within a Bloom filter (BF) that is transmitted along with the data. On receiving the packet, Base Station (BS) extracts the provenance information and verifies. The Extended provenance encoding scheme allows the Base Station to detect if packet drop attack was staged by consecutive malicious node. In existing system, a separate channels is used for the transmission of both data and provenance [4]. The traditional provenance security solutions employ append-based data structures to store provenance, leading to excessive cost as the size of the provenance data increases as the size of the network increases. Instead, the proposed scheme uses only fast Message Authentication Code (MAC) schemes and Bloom filters, where the size of data structures is fixed and compactly represents the provenance. Bloom filter yields low error rates in practice and do not require large bandwidth.

II. System Model

A. Network Model

Proposed scheme considers a multi-hop wireless sensor network that consists of a source node, number of intermediate sensor nodes and a base station which receives and collects the data packets transmitted over the network. The network is modeled as an acyclic graph. Each node in the network has a wireless link with other nodes in the network. The pair of nodes that are communicating directly has a distance which is taken as weight. The sensor nodes are stationary after deployment. The routing path may change over time due to node failure. Each node reports its neighboring node information to the base station after deployment. The base station assigns each node in the network a unique identifier, *Node-ID*. An AES key and a set of 3 Hash keys are distributed to each node in the network. These keys are used during provenance encoding.

B. Provenance Model

A node level provenance is considered which is encoded at each node to represent the presence of the specific packet. This helps in detecting selective forwarding attacks. Given packet *d* and its provenance data in an acyclic graph, of the network structure, where each vertex (*v*) in the graph is the intermediate node. Each vertex in the provenance graph is uniquely identified by a vertex ID (*VID*), that is generated by the base station. The Edge set *E* consists of directed edges with a distance parameter as weight connecting the consecutive nodes in the acyclic graph. Each data packet contains unique packet sequence number, data value, and provenance data. All nodes use the same path sequence number.

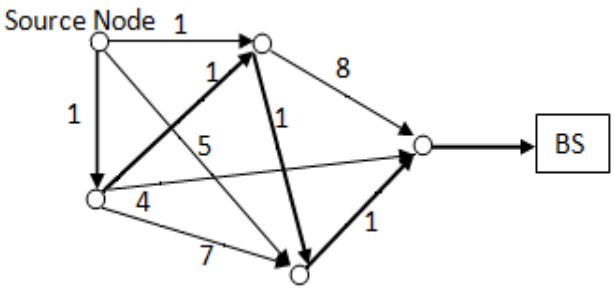

Fig:1. Provenance Graph for Sensor Network

C. Threat and Security

Base station is assumed to be trusted but, any intermediate nodes may be malicious. The adversary may perform traffic analysis, may deploy a few malicious nodes, or capture and compromise any existing node in the network and modify the memory contents. Denial of service attack is not considered as it makes the attack obvious. If the network node is compromised, the adversary may extract all key information. The adversary may drop, inject or alter data packets on the link under his control. Complete removal of provenance makes the data suspicious and hence, Base Station will be alarmed. So, the main concern is regarding misrepresentation of provenance data.

The adversary cannot obtain the knowledge of provenance data by analyzing the contents of the packets. Only base station is capable of analyzing the provenance data. This ensures confidentiality.

The adversary cannot add or remove data from provenance regarding any particular node in the network as the provenance data is represented using Bloom filter. Thus integrity is assured.

The adversary cannot replay the captured data as the base station can detect it due to usage of 3 Hash keys to encode Bloom Filter bits.

D. Bloom filter

The Bloom filter is a probabilistic data structure and is also space efficient [6]. It represents a set of items present in the set using an array of *m* bits and *k* Hash keys. Initially all bits in the Bloom Filter will be set to *0*. The result generated by each Hash key is used to map the presence of an item in the set, in the *m* bits of Bloom Filter. Every Item in the set will be inserted into the Bloom Filter by Hashing it with *k* Hash keys and resulting bits are set to *1* in the Bloom Filter. Bloom Filter allows False positive but not False negative; which means an element can be inserted as a member of the set, but cannot remove it once inserted without being detected.

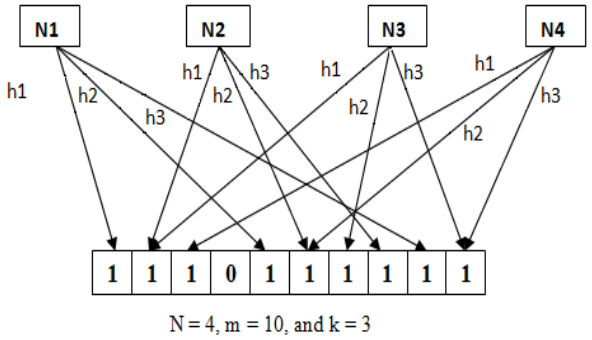

Fig: 2. Bloom Filter

III. Provinance Encoding

The proposed scheme uses a distributed mechanism for encoding and centralized mechanism for decoding provenance data at base station. Each received packet at base station consists of unique sequence number, data and in-packet Bloom filter.

Provenance encoding refers to generating packet sequence-ID Hashed value at each node using Hash keys and setting the respective bits to *1* in Bloom Filter. Each vertex is the node in the packet transmission path represented in the path sequence (*pseq*). VID is generates per packet based on packet sequence number. Thus for a given data packet, VID represents the specific node in the sequence path.

$$VID = generateVID\ (n_i, seq) = Ek_i\ (seq) \quad (1)$$

Where E is secure block cipher such as AES.

When the source node generates a packet, it also creates a iBF0, all bits initialized to zero. The source then generates a VID according to Eq. (1). Sets respective iBF0 bits to 1 and transmits the Bloom Filter as a part of the packet.

On receiving a packet, the intermediate nodes performs provenance aggregation as in Eq. (1) and forwards the packet to the next node in the path sequence.

IV. Provinance Decoding at Base Station

In Base Station, on receiving the packet it extracts the provenance data and does provenance verification. Base Station checks the iBF to see whether the path taken by the packet is same as the path it assumes given that all intermediate nodes in the path are active. If not, the provenance collection becomes necessary which retrieves the path traversed by the packet from the provenance collected and checks if it is a valid path. Invalid path includes illegitimate node in the path sequence. Using the path obtained, the base Station verifies the knowledge of provenance encoded in the packet.

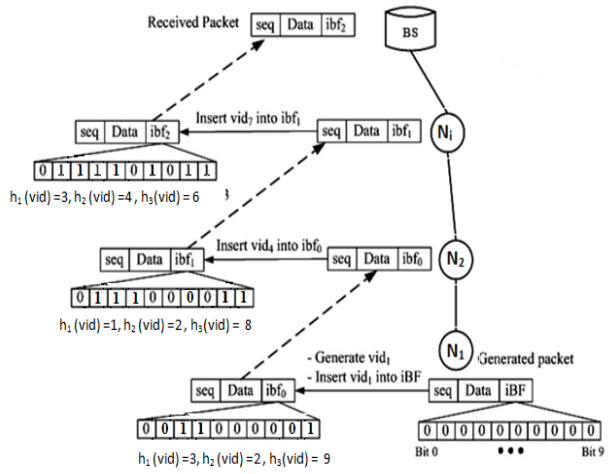

Fig: 3. Provenance encoding technique

The Base Station verifies the Provenance for both knowledge of the provenance as well as for the integrity of the transmitted provenance. The provenance verification succeeds only if the extracted iBF and generated BFc are equal. If not, the provenance collection process is triggered and retrieves the nodes traversed from the encoded provenance. Thus differentiates between the event of path change and the attack.

The second algorithm explains the provenance collection at the base station and the detection of the attack in multiple consecutive intermediate nodes.

An additional record can also be maintained to record the frequency of each provenance bit set to one and the intermediate node setting the respective bit to 1, to detect the number of packet drop in each intermediate malicious node in case of multiple consecutive malicious node.

A. *Algorithm 1 Provenance Verification*

Input: Received packet consisting packet sequence, path sequence and iBF.

Set of Hash Keys H, Data path $P = \{n1, ni, ..., np\}$

$BFc \leftarrow 0$ // Initialize Bloom Filter

for each n_i in P do

vid_i = generateVID (n_i,seq)

insert vid_i into BFc Using Hash function in H

endfor

if ($BFc = iBF$) **then**

return true // Valid Provenance

endif

return false

B. *Algorithm 2 Provenance Collection and Attack Detection*

Input: Received packet consisting packet sequence, path sequence and iBF.

N = the set of nodes in the network, H = set of Hash keys

Initialize S ← Null // set of possible nodes

Bloom Filter BFc ← 0

Determine possible nodes in the path and reconstruct the BF

for each node i N **do**

vid_i = genetateVID (i, seq)

if (vid_i is in iBF) **then**

S ← S U i

Set vid_i into BFc using Hash keys H

endif

endfor

Compare BFc and the iBF

If (BFc = iBF) **then**

Return S // Provenance is verified

else

Attacked Node AN = Null

for each node n_i in S **do**

vid_i = genetateVID (n_i, seq)

generate vid_i using Hash keys H

if (bits are not set to 1) **then**

AN ← AN U n_i

endif

endfor

endif

C. *Detection of paket drop Attack and Results*

The Provenance data obtained is used for malicious node detection, new packets inserted, number of packets dropped, number of packet dropped per malicious node.

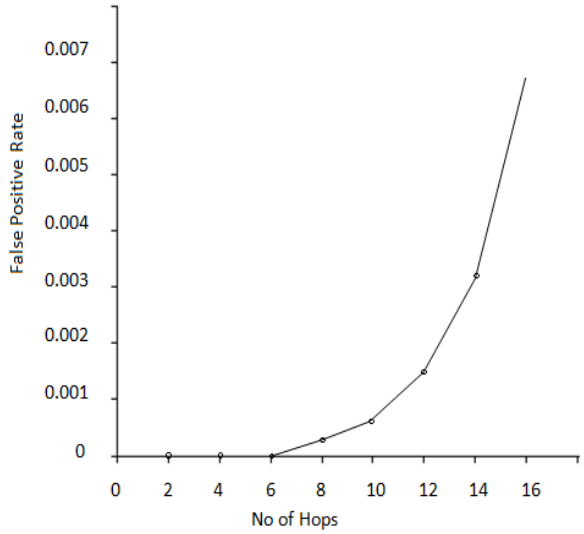

Fig: 4. False positive rate in provenance data

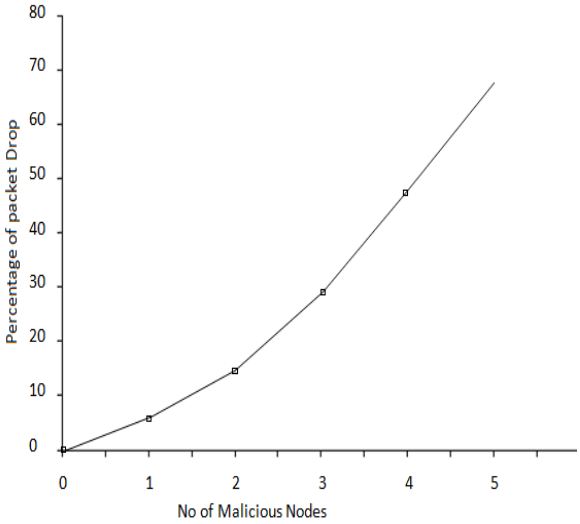

Fig: 5. Rate of Packet Drop for Multiple malicious consecutive nodes

Increase in the number of malicious node increases the overall percent of packet drop in the network. The percent of packet drop increases as the number of malicious nodes increases due to the decrease in the number of packets forwarded from the previous malicious nodes. Overall percent of packet drop per malicious node decreases as the number of packets forwarded is decreased after each malicious node is encountered.

V. CONCLUSION

The problem of secure transmission of provenance data in wireless sensor network is addressed and a light weight mechanism for provenance encoding and decoding based on Bloom Filter mechanism is proposed. The mechanism ensures the confidentiality and integrity of the Provenance data bound along with the actual data. The scheme detects the forgery of data packet and the packet drop attack in case of multiple consecutive malicious nodes. The result shows the effective detection of packet drop. In future, we plan on accurate forgery detection in case of multiple consecutive malicious nodes.

References

[1] Salmin Sultana, Gabriel Ghinita, Elisa Bertino and Mohamed Shehab, "Alightweight Secure Scheme for Detecting Provenance Forgery and Packet Drop Attacks in Wireles Sensor Networks," IEEE Transactions on Dependable and Secure computing, vol.12, No.3, May/June 2015.

[2] Vikash Kumar, Anshu Jain, and P.N.Barwal, "Wireless Sensor Networks: Security Issues, Challenges and Solutions," International Journal of Information and Computing Technology, ISSN 0974-2239 Volume 4, Number 8 (2014), pp. 859-868.

[3] Hani Alzaid, Ernest Foo and Juan Gonzalez Nieto, "Secure Data Aggregation in Wireless Sensor Network: A Survey," Information Security Institute, Qeensland University of Technology, PO Box 2434, Brisbane, Queensland 4001.

[4] GowriShankar.S, T.G. Bsavaraju, Manjaiah D.H, and Subir Kumar Sarkar, "Issues in Wireless Sensor Networks," Proceedings of the World Congress on Engineering 2008 Vol I, WCE 2008, July 2- 4, 2008, London, U.K.

[5] Salmin Sultana, Elisa Bertino, Mohamed Shehab, "A Provenance Based mechanism to Identify Malicious Packet Dropping Adversaries in Sensor Networks," Proceedings of the 2011 International Conferrence on Distributed Computing Systems Workshops, pages.332-338, 2011.

[6] Adam kirsch, Michael Mitzenmacher, "Less Hashing, Same Performance: Building a better Bloom Filter," accepted 11 july 2007, Published online 15 May 2008 in Wiley Inter Science.