IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555
Vol.6, No.2, Mar-April 2016

A Survey on Detection of Packet Drop Attack and Data Forgery using Dictionary Based Provenance in WSN

Sindhu J
Department of Computer Science and Engineering
B.M.S. College of Engineering
Karnataka, India

Mamatha P
Department of Computer Science and Engineering
B.M.S. College of Engineering
Karnataka, India

Kotramma Mathada
Department of Computer Science and Engineering
B.M.S. College of Engineering
Karnataka, India

Latha N.R, Assistant Professor
Department of Computer Science and Engineering
B.M.S. College of Engineering
Karnataka, India

----------***----------

Abstract--Data that originates in wireless sensor network (WSN) is processed by multiple intermediate processing nodes and traverse towards Base Station (BS). These sensor nodes often operate in an un-trusted environment where, adversary may introduce few malicious nodes in the network or compromises with the existing ones. Hence, it is necessary to address security requirements such as confidentiality, integrity and originality of data provenance. Data provenance allows the BS to trace the source and the forwarding path of an individual data packet. This paper is a brief survey on node level data provenance (which encodes history of data at each node) and the different techniques in WSN like in-packet-bloom filter, arithmetic coding and dictionary based provenance that make use of a light weight provenance scheme for detecting data forgery and packet drop attack. The paper also includes a brief survey on the use of provenance in WSN to overcome the above mentioned two major security attacks. In the In-packet boom filter technique, data provenance plays an important role for assuring data trustworthiness. But, the size of the provenance tends to increase with respect to increasing number of nodes in network. To overcome this Lossless arithmetic coding based compression technique is used to decrease the provenance size. In this compression technique, the provenance size is not directly proportional to the number of hops, but to the occurrence probabilities of the nodes that are on a packet's path. To make provenance size completely independent of number of nodes in a network a technique called dictionary based provenance is introduced. In this approach, each sensor node in the network stores a packet path dictionary. With the support of this dictionary, a path index instead of the path itself is enclosed with each packet. Since the packet path index is a code word of a dictionary, its size is independent of the number of nodes present in the packet's path.

Key Words: *Provenance, In-packet bloom filter, Dictionary provenance, Arithmetic coding*

1. INTRODUCTION

Wireless sensor networks are usually composed of hundreds and thousands of devices which are inexpensive but low-powered sensing devices with limited memory, computational and communication resources and limited batteries. WSN provides low-cost solutions in the applications such as battle-field surveillance, target tracking, environmental and health care monitoring, wildfire detection, traffic regulation etc. Sensor nodes comprise of a simple hardware that results in low deployment cost of WSN, but suffers from severe resource constraints. Thus the networks are vulnerable to many types of security attacks such as packet drop attack, false data injection and data forgery and eavesdropping.

This paper is a brief survey on two of the major security attacks in WSN Packet drop attack and data forgery. Packet drop is difficult to handle as it is not easy to determine whether it is a packet loss (i.e. due to signal problem or overload in network) or packet drop (i.e. Packet dropped by non-malicious node unintentionally which when detected is false data) or packet drop attack (i.e. Selective packet drop attack by malicious node which causes messages not to reach the intended destination hence required action may not happen). Data forgery detection is necessary as the data that traverses in WSN contain crucial information which is used in taking accurate decisions. Hence assuring trustworthiness of data is necessary that checks whether the data has been modified along its path which may result Base Station to take wrong decisions.

2. PROVENANCE

Generally Provenance is defined as a technique used to trace out the history of an object. Provenance records the history of data such as its place of origin, creator/publisher of data, creation date, modifier or modification date etc. The definition of provenance varies according to the application domains. In WSN provenance includes the origin of data packet, how it is processed through multiple intermediate nodes and its traversal in a network to reach the destination BS.

2.1 PROVENANCE IN WSN

In WSN provenance holds the history of data such as source node information and complete route information. Source node information includes information as to from where the packet originated and complete route information includes the path the data packet traversed form source node to BS. Each data packet contains: (i) a unique sequence number (ii) source node information (iii) data value (iv) provenance and (v) message authentication code (MAC).

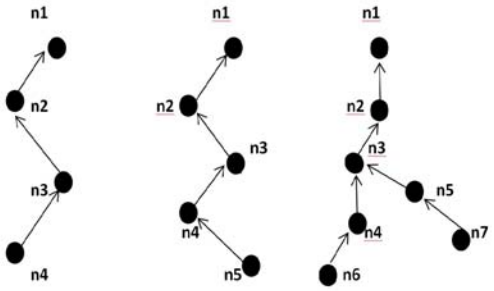

Fig -1: Provenance graphs of sensor networks

According to Changda Wang[20] provenance is stated as: Given a data packet d, the provenance p_d is a directed acyclic graph G(V,E) satisfying the following properties:

- P_d is a subgraph of the sensor network G(N,L).
- for v_x,v_y in V , v_y is a child of v_x if and only if HOST(v_y) forwards d to HOST(v_x).

Where V-Vertex, E-Edges , N-Network, L-Link .

Provenance transmission in Wireless Sensor Networks supports data transmission with non-negligible energy usage. In a multi-hop network, provenance includes knowledge of the originator and processing path of data since its generation. Thus, every intermediate node carries provenance of length proportional to the hop count between that node and the originator of the data item.

Consequently provenance information becomes complex and requires a large and variable number of bits in each packet which results in high energy dissipation. Chuang Wang et al[20] Proposes an energy-efficient provenance encoding and construction scheme known as Probabilistic Provenance Flow (PPF). The paper also demonstrates the feasibility of adapting the Probabilistic Packet Marking (PPM) technique in IP trace back in wireless sensor networks

Provenance management for streaming data requires addressing several challenges, including the assurance of high processing throughput, low bandwidth consumption, storage efficiency and secure transmission. Mohamed Shehab et al[19] discusses a novel approach to securely transmit provenance for streaming data by embedding provenance into the inter-packet timing domain. The challenge here is, with the increasing size of provenance, it should still be able to effectively manage and minimize the additional bandwidth consumption.

3. PACKET DROP ATTACK

Detecting Malicious Packet Losses is a challenging work since the normal network congestion can also produces the same effect. Modern networks usually drop packets when the load temporarily exceeds their buffering capacities. Some of the existing detection protocols have tried to address this problem with a user-defined threshold. One of the techniques uses a compromised router detection protocol that measures the traffic rates, buffer sizes and the number of congestive packet losses that occur. Using this information, the ambiguity is removed and subsequent packet losses are detected as malicious packet loss.

Taiwan et al[1] discusses the various negative impacts of packet dropping attacks that are as mentioned:

- Delay: The retransmissions of dropped packets in a FTP connection will drastically increase the total file transfer time.
- Response time: If the DNS query packets are dropped, a user may feel waiting for a long time to get a web page.
- Quality: Dropping some packets of MPEG video stream or IP telephoning data flow can degrade the quality of the service.
- Bandwidth: Because dropping packets usually introduces packet retransmissions, which leads to wastage of network bandwidth.

3.1 PACKET DROPPING ATTACK DETECTION TECHNIQUES

Kennedy Edemacu et al [2] propose several techniques to deal with the packet drop attack:

(i)Watch Dog Technique: In this technique, every node acts as a watchdog agent monitoring packet transmissions to neighboring nodes. The watchdog agents save a copy of packets in their watchdog monitoring buffers before they are transmitted to the next node. This helps to monitor the packet relay from a neighboring node to the next node.

(ii)Side Channel Monitoring (SCM) : In SCM a sub-set of neighbors for each node in a route between source and destination are selected to observe and monitor their message forwarding behaviors

(iii)Monitoring Agent Technique: The technique is based on capturing packets sent by neighboring nodes within a transmission range. All the nodes in a network collect information about their one hop neighbors within a certain period of time

(iv)PathRater: In this technique a PathRater is run by every node in the network. A node maintains ratings for every other node it knows in the network based on the knowledge of misbehaving nodes and link reliability of data in order to choose the most suitable path. A path metric is calculated by averaging the ratings for nodes in the path

According to V. Bhuse et al[3] the existing techniques to detect packet drop attack needs continuous monitoring of every node in the network. Once malicious nodes that drop packets are detected, a new path has to be found that does not include them. The paper proposes a lightweight solution

called DPDSN which identifies the paths that drop packets by using alternate paths, these alternate paths are the paths that WSN finds earlier during route discovery. According to DPDSN alternate path should not contain any node in common. There will be multiple path available from source to destination and the source sends the same message to the destination using multiple path. The source in turn asks destination to send back value of N i.e number of packets it has received from multiple path. Once the source receives the number of packets delivered to the destination from Multiple path, source is able to identify the path where packet-drop-attack has occurred based on the N value and also identifies the path that is safe for further communication.

Chuang Wang et al[4] proposes a simple yet effective scheme called as Probabilistic Nested Marking (PNM) that recognizes the misbehaving forwarders who drop or modify packets. The scheme identifies packet modifiers with a certain probability. In PNM, a dynamic routing tree rooted at the BS is first established, when sensor data is transmitted along the tree structure towards the BS, each packet sender or forwarder adds a small number of extra bits, which is called packet marks, to the packet. The format of the small packet marks is designed such that the BS can figure out the dropping rate associated with every sensor node, and then run the node categorization algorithm to identify the nodes that are droppers/modifiers or suspicious droppers/modifiers. Once the information of node behaviors has been accumulated, the BS periodically runs the heuristic ranking algorithm to identify the most likely bad nodes from the suspiciously bad nodes.

N. Vanitha et al[5] proposes a node categorization algorithm that recognizes which sensor node is the actual packet dropper. The algorithm also distinguishes the actual packet dropper from suspicious packet dropper. Identifying the actual packet dropper is a 3-step process that includes:

(1)Initialization Phase : During this phase the sensor nodes form topology of ToD(Tree on DAG) form

(2)In each round the data is transferred through routing tree from source node to BS. Sender/forwarder add packet marks to each packet. After completion of one round of transmission based on packet marks received by BS the node categorization algorithm is applied to find nodes that are bad for sure (i.e., packet droppers), suspiciously bad (i.e., suspected to be packet droppers) and good for sure (i.e., no packet droppers).

(3)In each round BS station receives the information using different routing topology. After certain amount of rounds passed with different topology BS receive information about the behavior of each sensor node in network, this information help in detecting packet dropper .

Salmin sultana et al[5] proposes that detection of malicious node involves 3 phases (i) Detecting Packet Loss (ii) Identification of Attack Presence (iii) Localizing the Malicious Node/Link.
The presence of packet loss is detected by checking inter-packet delays. The packet drop attack is determined by comparing the observed average packet loss rate with the natural packet loss rate of the data flow path. To determine a malicious node, other than Node-ID, time stamp, hash value of the data etc. are added to the provenance. Once presence of attack is detected, BS trace backs the path and notifies the source and intermediate nodes in that path while receiving data packet in next round. BS requests each sending or forwarding node to add complete provenance of the last packet which it has received. Thus BS traces out the path with the help of provenance and detects the malicious node that drop packet in that path.

4. DATA FORGERY DETECTION

Any malicious sensor node in network can inject false data during both data aggregation and data forwarding. Some of the existing techniques prevent false data injections during data forwarding by not allowing the forwarding node to modify the data packet.
Garaga Subba Rao et al[8] proposes a technique that prevent false data injections during both data forwarding and data aggregation using DSP(Dynamic Security Protocol). Traditional symmetric key cryptography algorithms, does not achieve both end-to-end confidentiality and network data aggregation.
This is achieved by dynamic security protocol that includes an efficient data aggregation algorithm where the messages are encrypted hop-by-hop. Thus in order to perform data aggregation, intermediate nodes have to decrypt each received message, then aggregate the messages according to the corresponding aggregation function, and finally encrypt the aggregation result before forwarding it . Thus, this is not an energy efficient way of performing secure data aggregation and it may result in considerable delay and the process also requires neighboring data aggregators to share secret keys for decryption and encryption.

In some of large-scale sensor network individual sensors are subject to security compromises. Compromised node can inject into the network large quantities of bogus sensing reports which can cause not only false alarms but also the depletion of the finite amount of energy in a battery powered network.

S. N. Saranya et al[7] proposes a Statistical En-route Filtering mechanism (SEF) technique that prevents any single compromised node from breaking down the entire system. This technique carefully limits the amount of security information assigned to any single node and collectively generates a legitimate report that carries multiple message authentication codes (MACs). The report with inadequate number of MACs will not be delivered. The sensing report are forwarded towards the BS over multiple

hops where each forwarding node verifies the correctness of the MACs carried in the report with certain probability. Once an incorrect MAC is detected, the report is dropped. The probability of detecting incorrect MACs increases with the number of hops the report travels. Depending on the path length, there is a non-zero probability that some reports with incorrect MACs may escape en-route filtering and be delivered to the BS. In any of the case the BS will further verify the correctness of each MAC carried in each report and reject false ones.

5. Packet Drop Attack And Forgery Detection Using Provenance

5.1 Bloom Filter

In-packet Bloom-filter[10] is one of the technique used for provenance encoding. Provenance encodes history of data at each node, therefore provenance size increases with the increase in the number of nodes in network. This is in-efficient as performance decreases due to high bandwidth consumption. Thus the main focus is to make provenance size light weight, secure transmission with forgery detection and finding packet drop attack using provenance data. This results in decreased bandwidth and energy which is the key factor in WSN. The second goal is to design a provenance encoding and decoding mechanism that satisfies security and performance needs by assuring confidentiality, integrity and originality of provenance .

Provenance size is defined by in-packet Bloom-filter size. Bloom filter is a fixed size data structure that is used to store provenance . Initially each bits in bloom filter are initialized to zero, when data flows through multi-hop, Vid(vertex id) is generated at each node. Hash function is applied on the generated Vid and then based on the output of hash function corresponding bits in bloom-filter is set to 1. This technique uses node level provenance which encodes the provenance at each node that are involved in each step of packet path. The resultant provenance size is directly proportional to number of nodes in network

5.2 Arithmetic Coding

Arithmetic coding is a technique used for data compression. Arithmetic coding can be used to reduce provenance size. This technique overcomes most of the disadvantages of previous techniques. Hussain et al[20] discusses that provenance size is not directly proportional to the number of hops, but to the occurrence probabilities of the nodes that are on a packet's path. Main focus of this technique is on compressing the provenance size because of bandwidth and energy limitation in WSN. Apart from this, there are other schemes for data compression that are based on Bloom filters or probabilistic packet marking approaches which have high error rates in provenance-recovery. However, some schemes drop critical information while compressing provenance record and do not include the edges that indicate directed connections among sensor nodes and thus fail to provide accurate packet path topologies.

Using distributed and lossless Arithmetic coding based compression technique results in a compression ratio higher than that of existing techniques. The Compression or encoding technique ensures that the system does not lose any provenance information after decoding.

5.3 Dictionary Based Provenance

Changda Wang et al[21] proposes a dictionary based provenance approach that makes provenance size completely independent of number nodes in the network. The provenance size is independent of hop count hence gives high utilization of bandwidth and less energy consumption. In dictionary based secure provenance scheme each sensor node in the network stores a packet path dictionary(PPD) . Using PPD, instead of entire path a path index is enclosed with each packet. Since the packet path index is a code word of a dictionary, its size is independent of the number of nodes present in the packet's path.

3. Conclusions

This paper is a brief survey on major security attacks and its impact on network. The paper discusses on two major security attacks packet drop attack and data forgery. Several existing technique to detect these two attacks and their disadvantage discussed. After a brief survey on provenance and its application in network, it is analyzed that use of light weight provenance scheme for detection of packet drop attack and data forgery in wireless sensor network yields better bandwidth utilization.

Acknowledgement

We would like to thank TEQIP-II (Technical Education Quality Improvement Programme) for giving us an opportunity to carry out our survey.

References

[1]. Xiaobing Zhang S. F. Wu ; Zhi Fu ; Tsung-Li Wu "Malicious Packet Dropping: How It Might Impact the TCP Performance and How We Can Detect It",pages.263-270,2000.

[2]. Kennedy Edemacu , Martin Euku and Richard Ssekibuule ,"Packet Drop Attack Detection Techniques In Wireless Ad Hoc Networks" ,vol.6,September 2014.

[3]. V. Bhuse, A. Gupta, and L. Lilien, "DPDSN: Detection of Packet-Dropping attacks for Wireless Sensor Networks"

[4]. Chuang Wang, Taiming Feng, Jinsook Kim, Guiling Wang, and Wensheng Zhang ," Catching Packet Droppers and Modifiers in Wireless Sensor Networks",vol.23,issue-5,pages.835-843,April 2011

[5]. N. Vanitha, G.Jenifa," Detection of Packet Droppers in Wireless Sensor Networks Using Node Categorization Algorithm".

[6]. Salmin sultana ,Elisa bertino, Mohamed Shehab "A Provenance based Mechanism to Identify Malicious Packet

Dropping Adversaries in Sensor Networks",pages.332-338,2011

[7]. M. Tharani, K. Sivachandran, S. N. Saranya, " An Efficient Detection Of Forgery And Packet Drop Attacks In Wireless Sensor Networks",vol.2,issue-7,Nov-2015.

[8]. Garaga Subba Rao, Kothapalli Ramesh, "False Data Detection in Wireless Network using Dynamic Security Protocol",vol.3,pages.4718-4722,2012.

[9]. S. Sultana, G. Ghinita, E. Bertino, and M. Shehab, "A lightweight secure provenance scheme for wireless sensor networks," in 2012 IEEE 18th International Conference on Parallel and Distributed Systems (ICPADS), 2012, pp. 101–108

[10]. S. Sultana, G. Ghinita, E. Bertino, and M. Shehab, "A lightweight secure scheme for detecting provenance forgery and packet drop attacks in wireless sensor networks," IEEE Transactions on Dependable and Secure Computing, vol. 99, no. PrePrints, p. 1, 2014.

[11]. I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," ACM, vol. 30, no. 6, pp. 520–540, 1987.

[12]. Hussain , Syed Rafiul , Wang, Changda, Sultana, Salmin , and Bertino, Elisa, "Secure Data Provenance Compression Using Arithmetic Coding in Wireless Sensor Networks" (2014).Cyber Center Publications.Paper 645.

[13]. Wang Changda , Hussain S and Bertino E "Dictionary based secure provenance compression for wireless sensor network ", ISSN:1045-9219, 2015 ,Volume:pp , Isssue:99.

[14]. H.S. Lim, Y.S. Moon, and E. Bertino, "Provenance-based trustworthiness assessment in sensor networks," in Proceedings of the Seventh International Workshop on Data Management for Sensor Networks, 2010, pp. 2–7.

[15]. Alam and S. Fahmy, "A practical approach for provenance transmission in wireless sensor networks," Ad Hoc Networks, vol. 16, no. 0, pp. 28 – 45, 2014.

[16]. E. Dawson, D. Wong, and D. Ma, "Secure feedback service in wireless sensor networks," in Information Security Practice and Experience. Springer Berlin Heidelberg, vol. 4464, pp. 116–128.

[17]. S. C. Misra, I. Woungang, S. Misra, A.-H. Jallad, and T. Vladimirova, "Data-centricity in wireless sensor networks," in Guide to Wireless Sensor Networks. Springer London, 2009, pp. 183–204.

[18]. B. Shebaro, S. Sultana, S. R. Gopavaram, and E. Bertino, "Demonstrating a lightweight data provenance for sensor networks," in ACM Conference on Computer and Communications Security, 2012, Conference Paper, pp. 1022–1024.

[19]. Salmin Sultana, Mohamed Shehab, and Elisa Bertino, "Secure Provenance Transmission for Streaming Data",vol.25,issue-8,pages.1890-1903,2013.

[20]. Syed Rafiul Hussain, Changda Wang, Salmin Sultana, and Elisa Bertino, "Secure Data Provenance Compression Using Arithmetic Coding in Wireless Sensor Networks",pages.645,Dec-2015

[21]. Changda Wang, Syed Rafiul , Hussain, and Elisa Bertino," Dictionary Based Secure Provenance Compression for Wireless Sensor Networks",vol.27,issue-2,pages.405-418,2015