Efficient Querying and Maintenance of Network to Identify Malicious Packet Dropping

M.Sindhu Lakshmi[1], T. Baba[2]

M.Tech, Dept. of CSE, P.V.K.K Institute of Technology, Ananthapuramu, Affiliated to JNTUA, India[1]

Assistant Professor, Dept. of CSE, P.V.K.K Institute of Technology, Ananthapuramu, Affiliated to JNTUA, India[2]

ABSTRACT: The numerous applications should work in Large-scale sensor networks domains. The data collected from wireless sensor network are used in making decisions in critical infrastructures. Data's are originated from multiple sources and transmitted through intermediate processing nodes. Those nodes perform the aggregation on information. An attacker compromise those type of networks by introducing additional nodes in the network or compromising the existing nodes. So achieving the high data trustworthiness is crucial for correct decision-making. While evaluating the trustworthiness of sensor data provenance is an important factor. The several challenging requirements for provenance management in sensor networks are low energy and low bandwidth consumption, competent storage and secure transmission. This survey proposes a new lightweight scheme in order to securely transmit provenance with sensor data. The proposed in-packet Bloom filters techniques used to encode provenance with the sensor data. This mechanism initially performs provenance at the base station then perform reconstruction of the data at the base station. In addition to this the provenance scheme functionality used to detect packet drop attacks organized by malicious data forwarding nodes. This survey describes the effectiveness and efficiency of the Light weight secure provenance scheme in detecting packet forgery and packet loss attacks..

I. INTRODUCTION

Sensor networks are becoming increasingly popular in numerous application domains, such as cyber physical infrastructure systems, environmental monitoring, power grids, etc. Data are produced at a large number of sensor node sources and processed in-network at intermediate hops on their way to a base station that performs decision-making. The diversity of data sources creates the need to assure the trustworthiness of data, such that only trustworthy information is considered in the decision process. Data provenance is an effective method to assess data trustworthiness, since it summarizes the history of ownership and the actions performed on the data. Recent

research highlighted the key contribution of provenance in systems where the use of untrustworthy data may lead to catastrophic failures e.g. SCADA systems for critical infrastructure. Although provenance modeling, collection, and querying have been investigated extensively for workflows and curated databases, provenance in sensor networks has not been properly addressed. In this paper, we investigate the problem of secure and efficient provenance transmission and processing for sensor networks. In a multi-hop sensor network, data provenance allows the base station to trace the source and forwarding path of an individual data packet since its generation. Provenance must be recorded for each data packet, but important challenges arise due to the tight storage, energy and bandwidth constraints of the sensor nodes. Therefore, it is necessary to devise a light-weight provenance solution which does not introduce significant overhead. Furthermore, sensors often operate in an untrusted environment, where they may be subject to attacks. Hence, it is necessary to address security requirements such as confidentiality, integrity and freshness of provenance. Our goal is to design a provenance encoding and decoding mechanism that satisfies such security and performance needs. We propose a provenance encoding strategy whereby each node on the path of a data packet securely embeds provenance information within a Bloom filter, which is transmitted along with the data. Upon receiving the data, the base station extracts and verifies the provenance. Sensor networks are used in numerous application domains, such as cyber physical infrastructure systems, environmental monitoring, power grids, etc. Data are produced at a large number of sensor node

sources and processed in-network at intermediate hops on their way to a Base Station (BS) that performs decision-making. The diversity of data sources creates the need to assure the trustworthiness of data, such that only trustworthy information is considered in the decision process. Data provenance is an effective method to assess data trustworthiness, since it summarizes the history of ownership and the actions performed on the data. Recent research highlighted the key contribution of provenance in systems where the use of untrustworthy data may lead to catastrophic failures (e.g., SCADA systems).Although provenance modeling, collection, and querying have been studied extensively for workflows and curated databases provenance in sensor networks has not been properly addressed. We investigate the problem of secure and efficient provenance transmission and processing for sensor networks, and we use provenance to detect packet loss attacks staged by malicious sensor nodes. In a multi-hop sensor network, data provenance allows the BS to trace the source and forwarding path of an individual data packet. Provenance must be recorded for each packet, but important challenges arise due to the tight storage, energy and bandwidth constraints of sensor nodes.

Therefore, it is necessary to devise a light-weight provenance solution with low overhead. Furthermore, sensors often operate in an untrusted environment, where they may be subject to attacks. Hence, it is necessary to address security requirements such as confidentiality, integrity and freshness of provenance. Our goal is to design a provenance encoding and decoding mechanism that satisfies such security and performance needs. We propose a provenance encoding strategy whereby each node on the path of a data packet securely embeds provenance information within a Bloom filter that is transmitted along with the data. Upon receiving the packet, the BS extracts and verifies the provenance information. We also devise an extension of the provenance encoding scheme that allows the BS to detect if a packet drop attack was staged by a malicious node. As opposed to existing research that employs separate transmission channels for data and provenance, we only require a single channel for both. Furthermore, traditional provenance security solutions use intensively cryptography and digital signatures, and they employ append-based data structures to store provenance, leading to prohibitive costs. In contrast, we use only fast Message Authentication Code (MAC) schemes and Bloom filters (BF), which are fixed-size data structures that compactly represent provenance. Bloom filters make efficient usage of bandwidth, and they yield low error rates in practice. Our specific contributions are:

I.   *We formulate the problem of secure provenance transmission in sensor networks, and identify the challenges specific to this context;*

J.   *We propose an in-packet Bloom filter provenance encoding scheme.*
K.   *We design efficient techniques for provenance decoding and verification at the base station;*
L.   *We extend the secure provenance encoding scheme and devise a mechanism that detects packet drop attacks staged by malicious forwarding sensor nodes.*

II. RELATED WORK

Ramachandran proposed Pedigree provenance scheme in which each packet is tagged with provenance data. Tagger is deployed at each host which tags each packet with provenance data. Paper used provenance data for traffic classification and Arbiter is deployed at each host which decides what to do with received packets having specific tags. Packet classification before Pedegree is mainly dependent on the IP addresses and port numbers but, after pedigree it has used tag information on the tags for packet classification. Pedegree scheme does not consider adversary network case and hence cannot deal with forgery attacks in the WSN.

Foster, J. Vockler addressed that network accountability and failure analysis is important for network management. It also described the need of network provenance. proposed ExSPAN provenance system in distributed environment. ExSPAN used data provenance to prove the state of the network. ExSPAN was developed using rapidnet which is based on ns3 toolkit. Experimental results showed that the system is generic and extensible. Same as Pedegree, this scheme also did not consider security of the provenance data.

W. Zhou, M. Sherr, T. Tao observed the need of securing the provenance information and proposed a scheme named, Secure Network provenance which gives proof for the state of the provenance data. Network operator can detect faulty nodes and also can assess the damage to network from such faulty nodes. Snoopy named SNP is proposed in paper and experimental results showed that Snoopy can prove state of provenance data in malicious WSN model. SNP scheme did not consider the limitations of WSN i.e. limited bandwidth, low battery and low memory.

K. Muniswamy-Reddy addressed the need to find source of the data which is transferred over the internet and proposed a Scheme which provides strong integrity and confidentiality of provenance data. Proposed scheme is designed in such way that it can be deployed at application layer Experiments showed that providing the integrity and Confidentiality to the provenance data results into overload with range 1% to 13%. Proposed approach gives control over the visibility of provenance data and assures no one can modify the provenance data without detection. Integrity and confidentiality is achieved through encryption and incremental chained signature mechanism.

Y. Simmhan, B. Plale, proposed a method to secure directed acyclic graph of the provenance data. Proposed method used digital signature in which provenance owner and processors tags or signs nodes. The relationship between provenance data graph and integrity is validated by checking the signatures. Both paper [4] and [5] are generic solutions which can be applied to any network and they are not designed with consideration of the nature of WSN Paper [6] proposed a mechanism in which sensor data is tagged with its provenance data automatically and provenance data can be recovered from this tagged data. Experiments with different scenarios proved robustness of
this scheme. Special feature of this scheme is that, the provenance data is embedded into actual sensor data. Proposed system does not provide any way to provide security to provenance data.

III.   EXISTING SYSTEM

Recent research highlighted the key contribution of provenance in systems where the use of untrustworthy data may lead to catastrophic failures (e.g., SCADA systems). Although provenance modeling, collection, and querying have been studied extensively for workflows and curated databases, provenance in sensor networks has not been properly addressed. Pedigree captures provenance for network packets in the form of per packet tags that store a history of all nodes and processes that manipulated the packet. Hasan et al. propose a chain model of provenance and ensure integrity and confidentiality through encryption, checksum and incremental chained signature mechanism.

IV. PROPOSED SYSTEM

We investigate the problem of secure and efficient provenance transmission and processing for sensor networks, and we use provenance to detect packet loss attacks staged by malicious sensor nodes. Our goal is to design a provenance encoding and decoding mechanism that satisfies such security and performance needs. We propose a provenance encoding strategy whereby each node on the path of a data packet securely embeds provenance information within a Bloom filter (BF) that is transmitted along with the data. Upon receiving the packet, the BS extracts and verifies the provenance information. We also devise an extension of the provenance encoding scheme that allows the BS to detect if a packet drop attack was staged by a malicious node.

A. *Data provenance at sensor network*
Sensor networks are used in various areas like such as cyber physical infrastructure systems, environmental weather monitoring, power grids, etc. Data are originated from a huge number of sensor node sources and they are processed at intermediate hops at in networks. These data's finally going to a base station (BS) which performs decision-making about where to go next. The uniformity of data sources creates assurance of the trustworthiness of data. This type of trustworthy information is considered in the decision making process at the base station. The data trustworthiness is assured by data provenance scheme. This is an effective method since it summarizes the history of ownership on the data and the list of actions performed on that information. The big advantage of this provenance scheme is detecting packet loss attacks organized by malicious/compromised sensor nodes. The major disadvantage of this scheme is the use of untrustworthy data at the nodes may create the catastrophic failures (e.g., SCADA systems). Although

provenance modeling, collection, and querying have been used extensively in workflows [1] and curated databases [2], provenance at sensor networks has not been fully addressed.

B. in packet Boom Filter(iBF)

This is a distributed mechanism in order to encode provenance at the nodes and it will work as centralized algorithm to decode it at the BS. The technical core of this survey is the notion of (iBF) [3]. In this packet consists of a unique sequence number, data value, and an iBF which contains the provenance. The focus of this scheme is a securely transmitting provenance with the data to the BS. In this aggregation framework, securing the data values is an important factor,. The secure provenance technique can be used to obtain a complete solution that provides security for data, provenance and data-provenance binding. The three Security Objectives in sensor networks is a confidentiality, Integrity and freshness.

C. Confidentiality

An attacker by analyzing the contents of a packet cannot gain any knowledge about data provenance. Only authorized users (e.g., the BS) can process the information and check the integrity of provenance.

D. Integrity

An attacker, acting individually or combining with others in a group, cannot add or remove non-colluding nodes. Also the attacker cannot add any data from the malicious user to the original data.

E. Freshness

An attacker cannot replay the captured data from the original user and ensure the provenance detected by the BS. It is also important to provide a coupling between data and provenance i.e. Data-Provenance Binding, so the attacker cannot successfully drop or alter the legitimate/valid data while containing the provenance with the data, or swapping the provenance of two packets.

F. Detecting Packet Drop Attacks

Provenance encoding could be used for a packet acknowledgement. By using this sensor can transmit more meta-data. For an any individual data packet, the provenance record generated by a node will now consist of the node ID and an acknowledgement in the form of a sequence number of the lastly seen (processed/forwarded) packet belonging to that data flow. If the intermediate packet could be drop by the attacker means some nodes on the path do not receive that packet. Hence, during the next round of packet transmission the mismatch between the acknowledgements should be generated from different nodes on the path. This factor could be to detect the packet drop attack and to localize the malicious node.
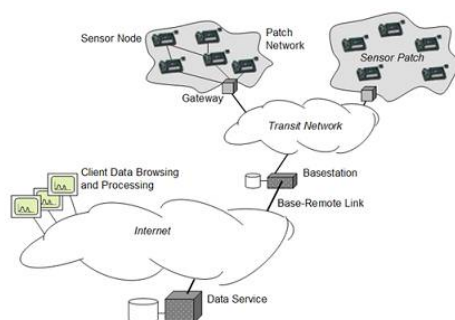
SYSTEM ARCHITECTURE

Fig 1 System Architecture

VI. CONCLUSION AND FUTURE WORK

We addressed the problem of securely transmitting provenance for sensor networks, and proposed a light-weight provenance encoding and decoding scheme based on Bloom filters. The scheme ensures confidentiality, integrity and freshness of provenance. We extended the scheme to incorporate data-provenance binding, and to include packet sequence information that supports detection of packet loss attacks. Experimental and analytical evaluation results show that the proposed scheme is effective, light-weight and scalable. In future work, we plan to implement a real system prototype of our secure provenance scheme, and to improve the accuracy of packet loss detection, especially in the case of multiple consecutive malicious sensor nodes.

REFERENCES

1. Foster, J. Vockler, M. Wilde, and Y. Zhao, "Chimera: A Virtual Data System for Representing, Querying, and Automating Data Derivation," Proc. Conf. Scientific and Statistical Database Management, pp. 37-46, 2002.
2. Muniswamy-Reddy, D. Holland, U. Braun, and M. Seltzer, "Provenance-Aware Storage systems," Proc. USENIX Ann. Technical Conf., pp. 4-4, 2006.
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. 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.
5. 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.
6. Y. Simmhan, B. Plale, and D. Gannon, "A Survey of Data Provenance in E-Science," ACMSIGMOD Record, vol. 34, pp. 31-36, 2005.
7. A. Ramachandran, K. Bhandankar, M. Tariq, and N. Feamster, "Packets with Provenance," Technical Report GT-CS-08-02, Georgia Tech, 2008.
8. W. Zhou, M. Sherr, T. Tao, X. Li, B. Loo, and Y. Mao, "Efficient Querying and Maintenance of Network Provenance at Internet-Scale," Proc. ACM SIGMOD Int'l Conf. Management of Data, pp. 615-626, 2010.
9. W. Zhou, Q. Fei, A. Narayan, A. Haeberlen, B. Loo, and M. Sherr, "Secure Network Provenance," Proc. ACMSOSP, pp. 295-310, 2011.
10. A. Syalim, T. Nishide, and K. Sakurai, "Preserving Integrity and Confidentiality of a Directed Acyclic Graph Model of Provenance," Proc. Working Conf. Data and Applications Security and Privacy, pp. 311-318, 2010.