ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

Provenance Forgery and Packet Drop Attacks Detection in Wireless Networks

Rasika Dinde[1], Apeksha Jain[2], Sachin Thorkar[3], Amarnath Patil[4]

Student, Computer, G.S.M.C.O.E, Savitribai Phule Pune University, Pune, India[123]

Professor, Computer, G.S.M.C.O.E, Savitribai Phule Pune University, Pune, India[4]

ABSTRACT: large scale sensor networks are deployed in numerous application domains, and the data they collect are used in decision-making for critical infrastructures. Data are streamed from multiple sources through intermediate processing nodes that aggregate information. A malicious adversary may introduce additional nodes in the network or compromise existing ones. Therefore, assuring high data trustworthiness is crucial for correct decision-making. Data provenance represents a key factor in evaluating the trustworthiness of sensor data. Provenance management for sensor networks introduces several challenging requirements, such as low energy and bandwidth consumption, efficient storage and secure transmission. In this paper, we propose a novel lightweight scheme to securely transmit provenance for sensor data. The proposed technique relies on in packet Bloom filters to encode provenance. We introduce 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 nodes. We evaluate the proposed technique both analytically and empirically, and the results prove the effectiveness and efficiency of the lightweight secure provenance scheme in detecting packet forgery and loss attacks.

KEYWORDS: Data provenance, Wireless Sensor Network, Bloom Filtering, Encryption, Decryption.

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, environmentalmonitoring, power grids, etc. Data are produced at a largenumber of sensor node sources and processed in-network atintermediate hops on their way to a Base Station (BS) thatperforms decision-making. The diversity of data sourcescreates the need to assure the trustworthiness of data,such that only trustworthy information is considered in thedecision process. Data provenance is an effective methodto assess data trustworthiness,

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

since it summarizes thehistory of ownership and the actions performed on thedata. Recent research [1] highlighted the key contribution ofprovenance in systems where the use of untrustworthy datamay lead to catastrophic failures (e.g., SCADA systems).Although provenance modeling, collection, and queryinghave been studied extensively for workflows and curateddatabases [2], [3], provenance in sensor networks has notbeen properly addressed. We investigate the problem ofsecure and efficient provenance transmission and processingfor sensor networks, and we use provenance to detect packetloss attacks staged by malicious sensor nodes. In a multi-hop sensor network, data provenance allowsthe BS to trace the source and forwarding path of anindividual data packet. Provenance must be recorded foreach packet, but important challenges arise due to the tightstorage, energy and bandwidth constraints of sensor nodes.

Therefore, it is necessary to devise a light-weight provenancesolution with low overhead. Furthermore, sensors oftenoperate in an untrusted environment, where they may besubject to attacks. Hence, it is necessary to address securityrequirements such as confidentiality, integrity and freshnessof provenance. Our goal is to design a provenance encodingand decoding mechanism that satisfies such security andperformance needs. We propose a provenance encodingstrategy whereby each node on the path of a data packetsecurely embeds provenance information within a Bloomfilter that is transmitted along with the data. Upon receivingthe packet, the BS extracts and verifies the provenanceinformation. We also devise an extension of the provenanceencoding scheme that allows the BS to detect if a packetdrop attack was staged by a malicious node.As opposed to existing research that employs separatetransmission channels for data and provenance [4], we onlyrequire a single channel for both. Furthermore, traditionalprovenance security solutions use intensively cryptographyand digital signatures [5], and they employ append-baseddata structures to store provenance, leading to prohibitivecosts. In contrast, we use only fast Message AuthenticationCode (MAC) schemes and Bloom filters (BF), which arefixed-size data structures that compactly represent provenance.Bloom filters make efficient usage of bandwidth,and they yield low error rates in practice. Our specificcontributions are:

1. ***We formulate the problem of secure provenance transmissionin sensor networks, and identify the challengesspecific to this context;***
2. ***We propose an in-packet Bloom filter provenance encoding scheme.***
3. ***We design efficient techniques for provenance decodingand verification at the base station;***
4. ***We extend the secure provenance encoding schemeand devise a mechanism that detects packet dropattacks staged by malicious forwarding sensor nodes;***
5. ***We perform a detailed security analysis and performanceevaluation of the proposed provenance encodingscheme and packet loss detection mechanism.The rest of the paper is organized as follows: Section 2 setsthe problem background and describes the system, threatand security models. Section 3 introduces the provenanceencoding scheme.***

II. RELATED WORK

Ramachandran [1] proposed Pedegree 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 [1] used provenance data fortraffic classification and Arbiter is deployed at each hostwhich 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.

Paper [2] addressed that network accountability and failure analysis is important for network management. It also described the need of network provenance. Paper [2] 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 [1] this scheme also did not consider security of the provenance data.

Wenchao Zhou [3] et.al 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

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

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.

Paper [4] 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.

Paper [5] 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.

Paper [7] focused on provenance management and proposed a novel secure provenance transmission schemein which provenance is embedded into inter packet timing domain and paper also considered limitations, requirements of WSN. Proposed scheme is different from traditional watermarking schemes. The scheme embeds provenance data into inter-packet delays and not in actual sensor data. As provenance data is not directly embedded into actual data, data quality degradation issue is solved. Provenance information is recovered using optimal threshold bases mechanism to reduce the provenance recovery errors. Proposed scheme is based on the spread spectrum watermarking technique and it is efficient against various sensor network or flow watermarking attacks. This scheme assumes that provenance dataremains same for flow of the packets.

Paper [8] described the design of the bloom filter data structure and its efficiency. Bloom filter is vector of n bits. When data is encoded into bloom filter, set of hash functions is used. Data to be encoded is hashed using hash function. Output of the hash function will be integer values. Initially bit vector contains all bit value equal to 0 bit. At output, integer index is set to 1. Main purpose of bloom filter is to check the membership of element i.e. once element is encoded; membership of the data can be checked. Paper [8] discussed the potential network applications of bloom filter data structure and described suitability of the bloom filter for network applications.

III. DETECTING PACKET DROP ATTACKS MECHANISM

We extend the secure provenance encoding scheme to detect packet drop attacks and to identify malicious node(s). We assume the links on the path exhibit natural packet loss and several adversarial nodes may exist on the path. For simplicity, we consider only linear data flow paths (i.e., as illustrated in Fig. 1(a)). Also, we do not address the issue of recovery once a malicious node is detected. Existing techniques that are orthogonal to our detection scheme can be used, which may initiate multipath routing [6] or build a dissemination tree around the compromised nodes [17]. We augment provenance encoding to use a packet acknowledgementthat requires the sensors to transmit more meta-data. For a 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 belongingto that data flow. If there is an intermediate packet drop,
Some nodes on the path do not receive the packet. Hence,during the next round of packet transmission, there will bea mismatch between the acknowledgements generated fromdifferent nodes on the path. We utilize this fact to detectthe packet drop attack and to localize the malicious node.We consider a data flow path P where nlis the only datasource. We denote the link between nodes niand n(i+1)as li. We describe next packet representation, provenanceencoding and decoding for detecting packet loss.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

1. Data Packet Representation

To enable packet loss detection, a packet header mustsecurely propagate the packet sequence number generatedby the data source in the previous round. In addition, as inthe basic scheme, the packet must be marked with a uniquesequence number to facilitate per-packet provenance generationand verification. Thus, in the extended provenancescheme, any jthdata packet contains (i) the unique packetsequence number (seq[j]), (ii) the previous packet sequencenumber (pSeq), (iii) a data value, and (iv) provenance.

2. Provenance Encoding

Fig. 4 depicts the extended provenance encoding process.The provenance record of a node includes (i) the nodeID, and (ii) an acknowledgement of the lastly observedpacket in the flow. The acknowledgement can be generatedin various ways to serve this purpose.

3. Provenance Decoding at the BS

Not only the intermediate nodes, but also the BS stores andupdates the latest packet sequence number for each dataflow. Upon receiving a packet, the BS retrieves the precedingpacket sequence (pSeq) transmitted by the source nodefrom the packet header, fetches the last packet sequence forthe flow from its local storage (pSeqb), and utilizes thesetwo sequences in the process of provenance verification andcollection.

IV. SYSTEM ARCHITECTURE

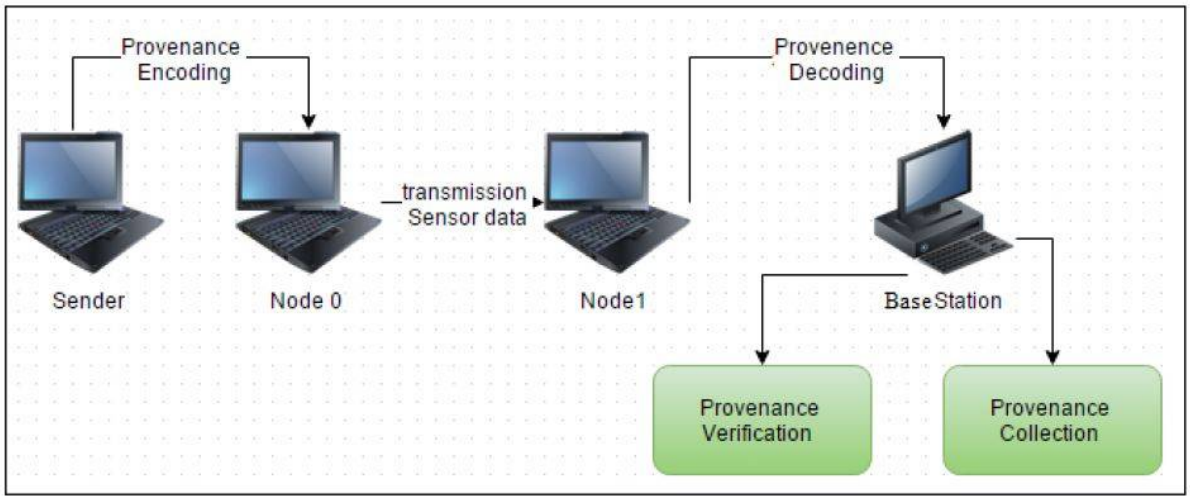


Fig No 01 System Architecture

Explanation-

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.

V. RESULTS

We implemented and tested the proposed techniques usingthe TinyOS simulator (TOSSIM) [24]. We have used themicazenergy model and PowerTOSSIM z [5] plugin toTOSSIM to measure the energy consumption. We considera network of 100 nodes and vary the network diameterfrom 2 to 14. All results are averaged over 100 runs. First,we look at how effective the secure provenance encodingscheme (introduced in Section 3) is in detecting provenanceforgery

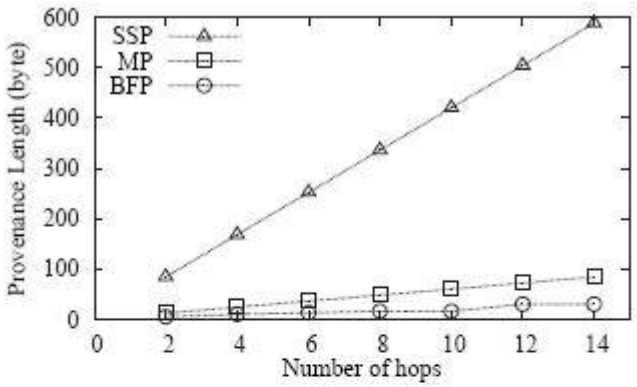
ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

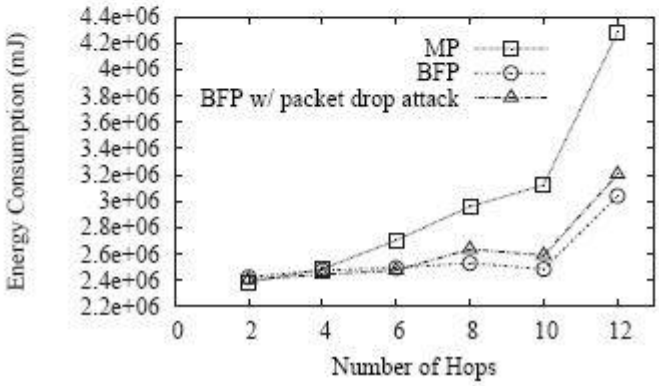
(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

and path changes. Next, we investigate the accuracyof the proposed method for detecting packet loss (whichwas presented in Section 4). Finally, we measure the energyconsumption overhead of securing provenance.



(a)



Compares SSP, MP and our provenance mechanismin terms of bytes required to transmit provenance. Theprovenance length in SSP and MP increases linearly withthe path length. For our scheme, we empiricallydeterminethe BF size which ensures no decoding error. Although theBF size increases with the expected number of elementsto be inserted, the increasing rate is not linear. We seethat even for a 14-hop path, a 30 byte BF is sufficient forprovenance decoding without any error.We also measure the energy consumption for both thebasic provenance scheme and the extended scheme forpacket drop detection, while varying hop counts. For packet drop attack, we set the malicious link loss rate as 0.03. Notethat, modern sensors use ZigBee specification for high levelcommunication protocols which allows up to 104 bytes asdata payload. Hence, SSP and MP can be used to embedprovenance (in data packet) for maximum 2 and 14nodes, respectively. Figure 8(b) shows aggregate energyconsumption over 1000 packet transmissions. The resultsconfirm the energy efficiency of our solutions.

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.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

REFERENCES

1. H. Lim, Y. Moon, and E. Bertino, "Provenance-based trustworthiness assessment in sensor networks," in Proc. of Data Management for Sensor Networks, 2010, pp. 2–7.
2. S. Sultana, E. Bertino, and M. Shehab, "A provenance based mechanism to identify malicious packet dropping adversaries in sensor networks," in Proc. of ICDCS Workshops, 2011, pp. 332–338.
3. L. Fan, P. Cao, J. Almeida, and A. Z. Broder, "Summary cache: a scalable wide-area web cache sharing protocol," IEEE/ACM Trans. Netw., vol. 8, no. 3, pp. 281–293, Jun. 2000.
4. Kirsch and M. Mitzenmacher, "Distance-sensitive bloom filters," in Proc. of the Workshop on Algorithm Engineering and Experiments, 2006, pp. 41–50.
5. 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.
6. S. Roy, M. Conti, S. Setia, and S. Jajodia, "Secure data aggregation in wireless sensor networks," IEEE Transactions on Information Forensics and Security, vol. 7, no. 3, pp. 1040–1052, 2012.
7. I. Foster, J. Vockler, M. Wilde, and Y. Zhao, "Chimera: A virtual data system for representing, querying, and automating data derivation," in Proc. of the Conf. on Scientific and Statistical Database Management, 2002, pp. 37–46.
8. K. Muniswamy-Reddy, D. Holland, U. Braun, and M. Seltzer, "Provenance-aware storage systems," in Proc. of the USENIX Annual Technical Conf., 2006, pp. 4–4.
9. Y. Simmhan, B. Plale, and D. Gannon, "A survey of data provenance in e-science," SIGMOD Record, vol. 34, pp. 31–36, 2005.
10. R. Hasan, R. Sion, and M. Winslett, "The case of the fake picasso: Preventing history forgery with secure provenance," in Proc. Of FAST, 2009, pp. 1–14.
11. S. Madden, J. Franklin, J. Hellerstein, and W. Hong, "TAG: a tiny aggregation service for ad-hoc sensor networks," SIGOPS Operating Systems Review, no. SI, Dec. 2002.
12. K. Dasgupta, K. Kalpakis, and P. Namjoshi, "An efficient clustering based heuristic for data gathering and aggregation in sensor networks," in Proc. of Wireless Communications and Networking Conference, 2003, pp. 1948–1953.
13. A. Kirsch and M. Mitzenmacher, "Distance-sensitive bloom filters," in Proc. of the Workshop on Algorithm Engineering and Experiments, 2006, pp. 41–50.
14. 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.
15. S. Roy, M. Conti, S. Setia, and S. Jajodia, "Secure data aggregation in wireless sensor networks," IEEE Transactions on Information Forensics and Security, vol. 7, no. 3, pp. 1040–1052, 2012.