ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

DISCERNING PROVENANCE FRAUDULENT USING SECURE SCHEME IN WSN

[1]Ms. R.VINOTHINI
Second Year [M.E]
Department of Computer Science and Engineering
Renganayagi Varatharaj College of Engineering
Sivakasi, Tamil Nadu, India
vinothinirajkumar4@gmail.com, Indian

[2]Mrs.D.KANAGAPUSHPAVALLI
Assistant Professor
Department of Computer Science and Engineering
Renganayagi Varatharaj College of Engineering
Sivakasi, Tamil Nadu, India
kanagadurai@gmail.com, Indian

Abstract- In wireless sensor network, sensors are used to cooperatively pass their data through the network to a main location. Based on continuous data flow from multiple sources and through intermediate processing by multiple aggregators the sensor networks are to be characterized. In sensor network, a black hole attack is a major security trouble to the data traffic, since it decreases the legitimate network output. That means, a harmful opponent may launch extra nodes in the network or damage existing ones. Low energy, Bandwidth consumption, efficient storage and secure transmission are called as challenging requirements introduced by provenance mechanisms. In this paper, we propose A Lightweight Scheme used to transmit sensor data provenance in secure manner. In provenance encoding, we use in-packet bloom filters. It is a space efficient data structure. In provenance verification at the base station we use provenance verification and collection algorithms. Finally, the results prove the power of Lightweight Secure Scheme for detecting provenance forgery in wireless sensor network.

Index terms – Provenance, security, sensor networks.

I. INTRODUCTION

Wireless Sensor Network is a self-forming network of small sensor nodes communicating among themselves using radio signals, and A deployed in quantity to sense, monitor and understand the physical world. The data gathered from wireless sensor network is usually rescue in the form of numerical data in a central base station. Additionally, the OGC is designate standards for interoperability interfaces and metadata encodings that enable real time unification of different sensor webs into the Internet, allowing any individual to monitor or control wireless sensor networks by a web browser. To reduce communication costs some algorithms remove or reduce nodes' expendable sensor information and avoid forwarding data that is of no use. As nodes can explore the data they forward, they can measure means or directionality for sample of readings from other nodes. For example, in sensing and watching applications, it is generally the case that proximating sensor nodes monitoring an environmental feature typically register similar values. This kind of data sacking due to the spatial correlation between sensor observations inspires techniques for in-network data aggregation and mining

Provenance helps gather, share and store the information which may lead to privacy and security concern in wireless sensor network. Security is one of the main characteristic of wireless sensor network affected with any attacks. Provenance, a mechanism of trust and reputation evaluation is an indispensable component to enhance the security of the entire network. Since provenance records the history of data acquisition and transmission, it is consideration as an effective mechanism to evaluate the trustworthiness and security of the data. It also provides the information about the operations performed on data. Reducing the size of the provenance is crucial in WSN as it is composed of a large number of sensor nodes. The limitation of provenance in WSN is tight storage, limited energy and increased bandwidth expenditure of the sensor node. Furthermore sensors often operate in an untrusted environment, where they may be subject to attacks. Provenance function is also deals with the detecting malicious node in network and to detect the packet drop in network. Provenance trustworthiness is very important in large scale sensor network as it is deployed in a military information network and trust assessment is a crucial task. In the computational world, as all kinds of information can easily be remaked, provenance becomes an important way of keeping track of alteration. Other applications of large scale network are medical monitoring, environmental monitoring, surveillance, home security, industrial machine monitoring etc.

Provenance management for sensor networks introduces several challenges such as low energy and bandwidth consumption, well planned storage and secure transmission. There are numerous techniques and method proposed for confidentiality, integrity, and trustworthy of secure provenance transmission in WSN. Distributed system which

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

evaluates the trust in the network that is more flexible and more responsive, which enhance the network trust in network. As trust is monitored and network is continuously restructured, our network remains trustworthy for a longer time.

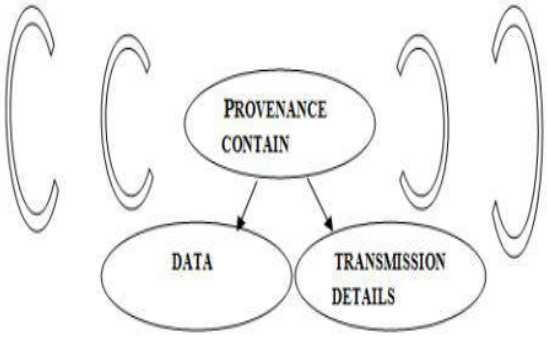

Figure 1 . Provenance model

Bloom filter is a manageable, space-efficient probabilistic data structure that succinctly represents a set in order to support membership queries. Bloom filter is implemented as a bit-array of m bits associated with h different hash functions, each of which maps a set constituent to one of the m array positions in a uniformly random manner. All bits in an initial Bloom filter are set to 0, position for an empty set. To insert an element u into a set represented by a Bloom filter BF, h array positions are deliberated by hash functions on u and the bits at those positions in BF are set to 1.It formulate the problem of secure provenance transmission in sensor networks.

II. RELATED WORK

[1]The key idea of digital watermarking is to hide a secret information (watermark) related to a digital content within the content itself thereby guarding the movement of the watermark along with the content. Thus, digital watermarking involves the collection of a watermark carrier domain and the design of two complementary Processes.[2] In the first step of this approach, the aggregator collects sensor data from nodes and computes the aggregation result. In the second step, the aggregator constructs a commitment based on Merkle hash-trees corresponding to the sensor data collected in the first step. In this construction, all the collected data is placed at the leaves of the tree, and the aggregator then evaluates a binary hash tree starting from the leaf nodes; each internal node in the hash tree is computed as the hash value of the sequence of the two child nodes. The root of the tree is called the commitment of the collected data. Because the hash function in use is collision resistant, once the aggregator commits to the collected values, it cannot alter any of the collected values. In the third step, the aggregator sends the home server both the aggregation result and its associated commitment. The home server and the aggregator engage in an interactive protocol in which the aggregator proves to the home server that the reported results are correct.[3] a novel and efficient scheme called SIES. SIES is the first solution that carries Secure In-network processing of Exact SUM queries, satisfying all security properties. It achieves this goal through a collaboration of homomorphic encryption and secret sharing. Furthermore, SIES is less weight (it relies on inexpensive hash operations and modular additions/multiplications), and features a very small bandwidth emaciation (in the order of a few bytes). Consequently, SIES constitutes an ideal method for resource-constrained sensors.[4] cyclic framework for computing trust scores. Through extensive experiments, we first show that our method works properly in sensor networks and the cyclic frame- work gradually evolves trust scores by reflecting changes in sensing value changes.[5] ExSPAN describes the history and derivations of network state that result from the execution of a distributed protocol. This system also does not address security concerns and is specific to some network use cases.[6] SNP expands network provenance to adversarial environments. Since all of these systems are general purpose network provenance systems, they are not optimized for the resource constrained sensor networks.

III.FLOW DESIGN

To encode provenance at the nodes and a centralized algorithm to decode it at the BS. The technical core of our proposal is the notion of in packet Bloom filter. Each packet consists of a unique sequence number, data value, and an iBF which holds the provenance. We emphasize that our focus is on securely transmitting provenance to the BS. In an aggregation infrastructure, securing the data values is also an important aspect, but that has been already addressed in previous work (e.g., [10]). Our secure provenance technique can be used in conjunction with such work to obtain a complete solution that provides security for data, provenance.

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

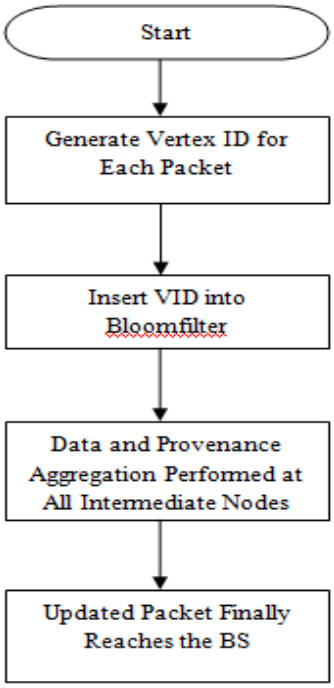

Figure 2 . Provenance Encoding

For a data packet, provenance encoding refers to generating the vertices in the provenance graph and inserting them into the iBF. Each vertex originates at a node in the data path and represents the provenance record of the host node. A vertex is uniquely identified by the vertex ID. The VID is generated per-packet based on the packet sequence number (seq) and the secret key Ki of the host node. We use a block cipher function to produce this VID in a secure manner.

When the BS receives a data packet, it executes the provenance verification process, which assumes that the BS knows what the data path should be, and checks the iBF to see whether the correct path has been followed. However, right after network deployment, as well as when the topology changes (e.g., due to node failure), the path of a packet sent by a source may not be known to the BS. In this case, a provenance collection process is necessary, which retrieves provenance from the received iBF and thus the BS learns the data path from a source node. Afterwards, upon receiving a packet, it is sufficient for the BS to verify its knowledge of provenance with that encoded in the packet.

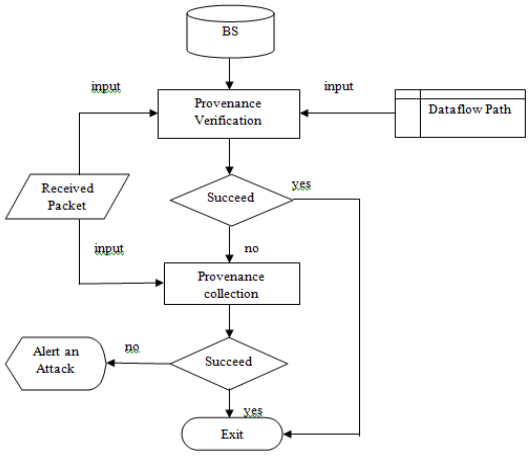

Figure 3 . Provenance decoding

A. NETWORK FORMATION

Wireless sensor network is created with number of sensor packet that collects data from the network. Each packet request generates data periodically and it aggregates to the base station. Data streamed from the multiple sources are aggregate in the intermediate processing nodes. The malicious adversary may introduce for the attacks.

We consider a multi hop wireless sensor network, consisting of a number of sensor nodes and a base station that collects data from the network. The network is modeled as a graph G(N,L). Sensor nodes are stationary after deployment, but routing paths may change over time, e.g., due to node failure. Each node reports its neighboring (i.e., onehop) node information to the BS after deployment. The BS assigns each node a unique identifier nodeID and a symmetric cryptographic key Ki. In addition, a set of hash functions H are broadcast to the nodes for use during provenance embedding.

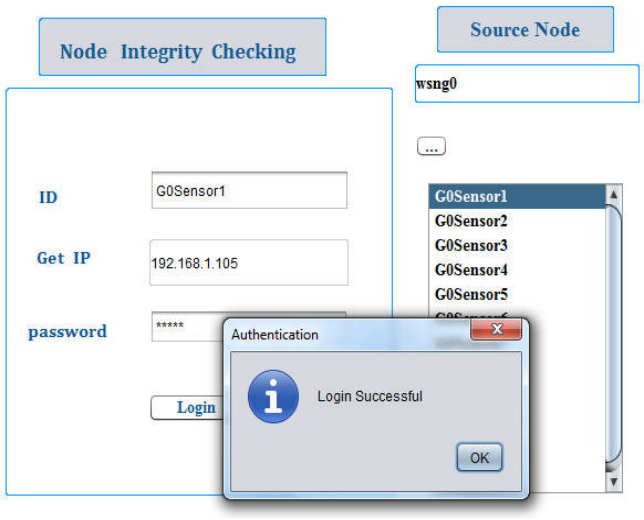

Figure 4 . Network formation

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

B. PROVENANCE MODELLING

Provenance summarizes the history of the ownership of items and the actions performed on them. Traditionally, people have used provenance to authenticate physical objects in arts, archives, and archaeology. But life today has become increasingly dependent on digital information that originated elsewhere, was processed by other people, and was stored in untrustworthy storage. So, it is increasingly important to know the source of the information and its derivation history. In other words, to be able to trust a piece of information, we need to know and verify its provenance. Data provenance has been defined in many ways, but all definitions share the same core concept: it is a description of the origin, derivation, and transmission history of data. Till now, scientists have been the primary users of data provenance systems. Provenance research has mainly focused on the tasks of modeling, collection, annotation, and querying. But as provenance steps into mainstream computing, new challenges arise. With its increased use in financial, medical, and other non-scientific application areas, provenance information faces a host of security threats, including active attacks from adversaries. In high-stakes business and medical applications, insiders may have significant incentives to alter data records' history. As information crosses application and organizational boundaries and passes through untrusted environments, its provenance becomes vulnerable to illicit alteration. When the trustworthiness of the provenance records themselves is in question: we need provenance of provenance, i.e., a model for secure provenance. Making provenance records trustworthy is challenging. Ideally, we need to guarantee completeness – all relevant actions pertaining to a piece of information are captured; integrity – adversaries cannot forge or alter provenance records; availability – auditors can verify the integrity of provenance information; confidentiality – only authorized parties can read provenance records;and efficiency – provenance mechanisms must have low overheads.

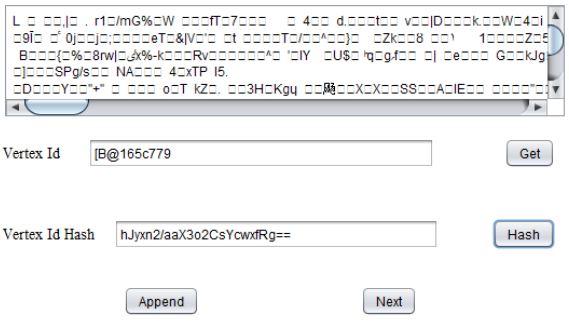

Figure 5 . Provenance modelling

C. DATA TRAVELLING

When a source node generates a packet, it also creates a BF (referred to as ibf0), initialized to 0. The source then generates a vertex. Inserts the VID into ibf0 and transmits the BF as a part of the packet. Upon receiving the packet, each intermediate node nj performs data as well as provenance aggregation. If nj receives data from a single child nj_1, it aggregates the partial provenance contained in the packet with its own provenance record. In this case, the iBF ibfj_1 belonging to the received packet represents a partial provenance, i.e., the provenance graph of the sub-path from the source up to nj_1. On the other hand, if nj has more than one child, it generates an aggregated provenance from its own provenance record and the partial provenance received from its child nodes. At first, nj computes a BF ibfj_1 by bitwise-ORing the iBFs from its children. ibfj_1 represents a partial aggregated provenance from all of the children. In either case, the ultimate aggregated provenance is generated by encoding the provenance record of nj into ibfj_1. To this end, nj creates a vertex and inserts the VID into ibfj_1 which is then referred to as ibfj. When the packet reaches the BS, the iBF contains the provenance records of all the nodes in the path i.e. the full provenance. We denote this final record by ibf. Example. The data path considered is <1; 4; 7>, where node 1 is the data source. We use a 10-bit BF and a set of three hash functionsH ={h1; h2; h3} for BF operations. When node 1 generates a data packet with sequence number seq, it creates the BF ibf0 which is set to all 0's. The node then creates a vertex corresponding to its provenance record and computes the VID as vid1 =EK1 (seq). To insert vid1 into ibf0, node 1 generates three indices as h1(vid1) = 1, h2(vid1) = 3, h3(vid1Þ= 8. The VID is then inserted by setting ibf0[1], ibf0[3], and ibf0[8] to 1. The updated ibf0 along with the packet is then sent towards the BS.

Upon receiving the packet, node 4 performs provenance aggregation. Since the node has one child, it only aggregates its own provenance record with ibf0. For this purpose, the node generates a VID vid4; computes 3 indices as h1(vid) = 3, h2(vid) = 6, h3(vid)= 9; and inserts vid4 into ibf0 by setting bits 3, 6, 9 of the iBF to 1. This updated iBF is referred to as ibf1. The data packet with ibf1 is then forwarded to node 7 which repeats the provenance aggregation. At the end, the BS receives the packet with the final Ibf (ibf2 from node 7) and stores this iBF for further processing.

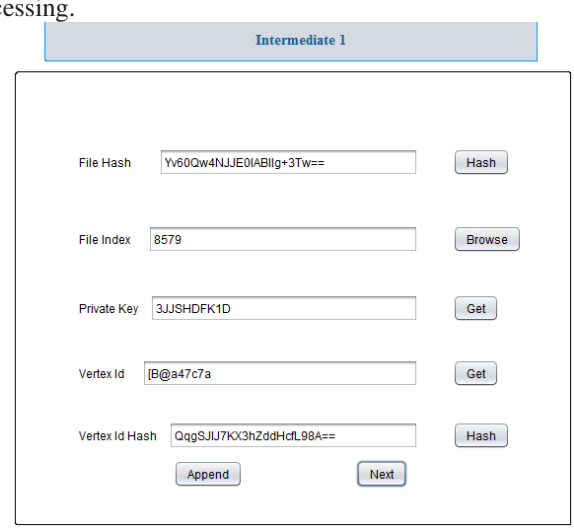

Figure 6 . Data travelling

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

D. VERIFICATION

The BS conducts the verification process not only to verify its knowledge of provenance but also to check the integrity of the transmitted provenance. We assume that the knowledge of the BS about this packet's path is P0. At first, the BS initializes a Bloom filter BFc with all 0's. The BF is then updated by generating the VID for each node in the path P0 and inserting this ID into the BF. BFc now reflects the perception of BS about the encoded provenance. To validate its perception, the BS then compares BFc to the received iBF ibf. The provenance verification succeeds only if BFc is equal to ibf. Otherwise, if BFc differs from the received iBF, it indicates either a change in the data flow path or a BF modification attack. The verification failure triggers the provenance collection process which attempts to retrieve the nodes from the encoded provenance and also to distinguish between the events of a path change and an attack.

provenance collection scheme makes a list of potential vertices in the provenance graph through the ibf membership testing over all the nodes. For each node ni in the network, the BS creates the corresponding vertex (i.e., vi with VID vidi) . The BS then performs the membership query of vidi within ibf. If the algorithm returns true, the vertex is very likely present in the provenance, i.e., the host node ni is in the data path. Once the BS finalizes the set of potential candidate nodes it executes the provenance verification algorithm on this set. This step is required to distinguish between the cases of a legitimate route change and that of malicious activity. If the verification succeeds, we decide that there was a natural change in the data path and we have been able to determine the path correctly. Otherwise, an attack has occurred.

The confidentiality of the scheme is achieved through two factors: the use of BF and the use of encryption keys. When one-way hash functions are used to insert elements in the BF, the identities of the inserted elements cannot be reconstructed from the BF representation. Attacker(s) may attempt to generate fake data and construct the provenance including some innocent nodes to make them responsible for false data and consequently to mark them as untrustworthy. However, the provenance embedding process requires the node specific secret Ki for cryptographic computation of the corresponding VID, and the attackers do not know the key for the legitimate nodes. Hence, this attack will fail.

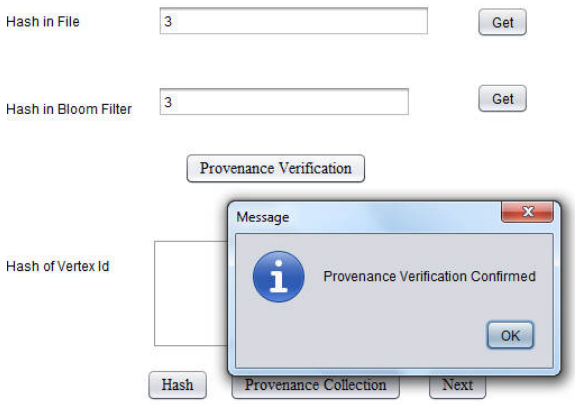

Figure 7 . Verification

III. 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] 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.

[2] 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.

[3] S. Sultana, E. Bertino, and M. Shehab, "A Provenance Based Mechanism to Identify Malicious Packet Dropping Adversaries in Sensor Networks," Proc. Int'l Conf. Distributed Computing Systems (ICDCS) Workshops, pp. 332-338, 2011.

[4] S. Papadopoulos, A. Kiayias, and D. Papadias, "Secure and Efficient In-Network Processing of Exact Sum Queries," Proc. Int'l Conf. Data Eng., pp. 517-528, 2011.

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

[6] 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.

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

[8] 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.

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

[10] P. Jokela, A. Zahemszky, C. Esteve, S. Arianfar, and P. Nikander, "LIPSIN: Line Speed Publish/Subscribe Inter-Networking," Proc. ACM SIGCOMM Conf. Data Comm., pp. 195-206, 2009.