Volume 3, Issue 4, April 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Confidential and Reliable Data Storage in WSN

Kusuma K V
Department of Computer Science & Engineering,
East West Institute of Technology,
VTU, India

Prasad M R
Assistant Professor
Department of Computer Science & Engineering,
East West Institute of Technology, VTU, India

Abstract: We consider two-tiered sensor networks, where data storage node serves as an intermediate tier between sensors and a sink for storing data and processing queries. This architecture has been widely adopted because of the benefits of power and storage saving for sensors as well as the efficiency of query processing. However, the importance of storage nodes also makes them attractive to attackers. In this paper, we propose a technique that prevents attackers from gaining information from sensor collected data. To preserve confidentiality we use encryption mechanism, so that data at the storage node is not available to attacker.

Keywords- Confidentiality, reliability, encryption, sensor networks.

I. Introduction

Wireless sensor network (WSN) is a heterogeneous system combining thousands to millions of tiny, inexpensive sensor nodes with several distinguishing characteristics. Wireless sensor networks (WSNs) have been widely deployed for various applications like security monitoring, environmental data collection, medical science, military, tracking etc. Security becomes extremely important factor when sensor networks are randomly deployed in a hostile environment. In this paper, we consider a two-tiered sensor network (shown in fig. 1) which consists of regular sensors and some special storage nodes, which are equipped with much larger storage than regular sensors. In this structure, regular sensors periodically forward the raw data to a nearby storage node and user queries are diffused to storage nodes by the sink. As an intermediate tier, storage nodes are responsible for hosting raw data and replying queries. For example, when deploying a sensor network in a building to monitor the environmental conditions, we may place a few storage nodes at each floor. In a habitat monitoring application, we may divide the wild filed into several regions and deploy one or a few storage node in each region.

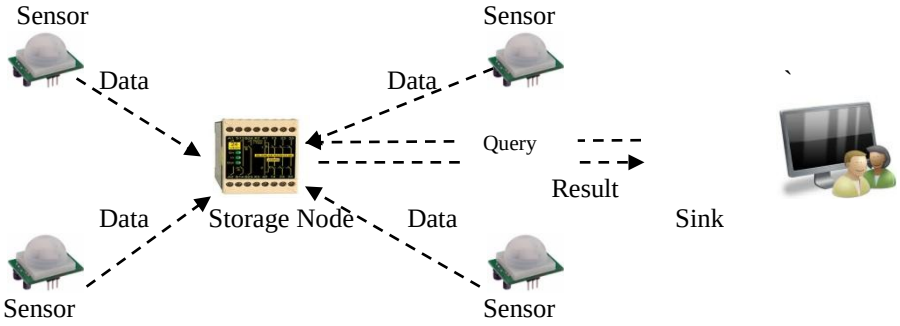

Fig.1 Architecture of two-tiered sensor networks

The storage nodes serve as an intermediate tier between the sensors and the sink for storing data and processing queries. Storage nodes bring three main benefits to sensor networks. First, sensors save power by sending all collected data to their closest storage node instead of sending them to the sink through long routes. Second, sensors can be memory-limited because data are mainly stored on storage nodes. Third, query processing becomes more efficient because the sink only communicates with storage nodes for queries. The inclusion of storage nodes in sensor networks was first introduced in [1] and has been widely adopted [2]-[6]. The inclusion of storage nodes also brings significant security challenges. As storage nodes store data received from sensors and serve as an important role for answering queries, they are more vulnerable to be compromised, especially in a hostile environment. A compromised storage node imposes significant threats to a sensor network. First, the attacker may obtain sensitive data that has been, or will be, stored in the storage node. Second, the compromised storage node may return forged data for a query. Third, this storage node may not include all data items that satisfy the query.

Therefore, we propose a scheme that prevents attackers from gaining information from sensor collected data and allows the sink to detect compromised storage nodes when they misbehave. For confidentiality, the storage node does not allow

unauthorized users to access information. For reliability, if the data is modified, sink is not provided the data until it is updated by the sensor. The main contributions of this paper include:

- We provide a novel mechanism for secure data storage in the encryption domain.
- We present an efficient data structure to guarantee the integrity of query results.

II. Related Work

A. Privacy and Integrity Preserving in WSNs

Privacy- and integrity-preserving range queries in WSNs have drawn people's attention recently [6], [7]. Sheng and Li proposed a scheme to preserve the privacy and integrity of range queries in sensor networks [6]. The basic idea is to divide the domain of data values into multiple buckets, the size of which is computed based on the distribution of data values and the location of sensors. In each time-slot, a sensor collects data items from the environment, places them into buckets, encrypts them together in each bucket, and then sends each encrypted bucket along with its bucket ID to a nearby storage node. For each bucket that has no data items, the sensor sends an encoding number, used by the sink to verify that the bucket is empty. When the sink wants to perform a range query, it finds the smallest set of bucket IDs that contains the range in the query, and then sends the set as the query to storage nodes. Upon receiving the bucket IDs, the storage node returns the corresponding encrypted data in those buckets. The sink can then decrypt the encrypted buckets, verify integrity using encoding numbers. Drawback of S&L scheme is that the bucket-partitioning technique allows compromised storage nodes to obtain a reasonable estimation on the actual value of data items and queries[8].

III. Models And Problem Statement

A. System Model

A Two-tiered sensor network consists of three types of nodes: sensors, storage nodes, and a sink. Sensors are inexpensive sensing devices with limited storage and computing power. They are often massively distributed in a field for collecting physical or environmental data, e.g., temperature. Storage nodes are powerful wireless devices that are equipped with much more storage capacity and computing power than sensors. Each sensor periodically sends collected data to its nearby storage node. The sink is the point of contact for users of the sensor network. Each time the sink receives a question from a user, it first translates the question into multiple queries and then disseminates the queries to the corresponding storage nodes, which process the queries based on their data and return the query results to the sink. The sink unifies the query results from multiple storage nodes into the final answer and sends it back to the user.

B. Threat Model

For a two-tiered sensor network, we assume that the sensors and the sink are trusted, but the storage nodes are not. In a hostile environment, both sensors and storage nodes can be compromised. If a sensor is compromised, the subsequent collected data of the sensor will be known to the attacker, and the compromised sensor may send forged data to its closest storage node. It is extremely difficult to prevent such attacks without the use of tamper-proof hardware. However, the data from one sensor constitute a small fraction of the collected data of the whole sensor network. Therefore, we mainly focus on the scenario where a storage node is compromised. Compromising a storage node can cause much greater damage to the sensor network than compromising a sensor. After a storage node is compromised, the large quantity of data stored on the node will be known to the attacker, and upon receiving a query from the sink, the compromised storage node may return a falsified result formed by including forged data or excluding legitimate data. Therefore, attackers are more motivated to compromise storage nodes.

C. Problem Statement

The fundamental problem for a two-tiered sensor network is the following: How can we design the storage scheme and the query protocol in a confidential and reliable manner. A satisfactory solution to this problem must meet the following two requirements.

1) Data confidentiality: The actual data stored in storage node can be viewed only by an authorized user. If an authorized user tries to view the data, sink can detect the misbehavior of storage node and unauthorized user is able to view irrelevant data.
2) Data reliability: If a query result that a storage node sends to the sink includes forged data or excludes legitimate data, the query result is not available to the sink until it is updated by the sensor.

IV. Modules Description

A. Sensor Module

Sensors send data to data storage node by splitting it into five blocks. For each block a unique HMAC code will be generated and stored in the database. Also a secret key will be generated for each file. If the data block stored in data storage node is modified, a new HMAC code will be generated for the modified data. In such case, the sensor can check if or not the data block is modified by comparing the HMAC code of the changed data with the HMAC code of the original data sent by the sensor. If the data block is modified, the sensor can update the modified block to original or it can delete the entire file.

B. Storage Node Module

Storage nodes are powerful wireless devices that are equipped with much more storage capacity and computing power than sensors. The storage node collects all data from the sensor nodes. The storage node allows only the Authorized user to

view the actual value of sensor node data. If any unauthorized user trying to view the sensor node data, he is able to view only encoded data.

C. Sink Module

The sink is the point of contact for users of the sensor network. Sink requests for secret key from sensor and then requests the file from the data storage node. If the file requested from sensor is unmodified and available, then sink receives the file. Sink can detect misbehavior of compromised storage nodes.

V. The Proposed Scheme

The data is sent to storage node by splitting it into five blocks. This reduces network bandwidth, network overhead, increases efficiency and reduces the chances of the attacker to understand the actual data even when the storage node is compromised.

Algorithm to provide confidentiality:

- Split the data into n data blocks $d_1, d_2, \ldots, d_n$ where n=5
- Encrypt the data using secret key $E(SK,d_i)$ where i=5
- Let each of the n data block contain m bytes of data $b_1, b_2, \ldots, b_m$ where m=1024
- For each data block sent, generate a hash code using SHA1, MD5

The data sent to storage node contains the following information:

Sensor ⟶ Storage node: $S_{id,}$ SK, $\{d1, d_2, \ldots, d_5\}$, HMAC $\{ d_1, d_2, \ldots, d_5\}$

Where, S_{id} is the Sensor ID, SK is the Secret Key.

Confidentiality in storage node is ensured by providing access only to authorized users through the secret key shared between sensor and storage node (as shown in fig 2).

Algorithm to provide reliability:

- The HMAC values of the data blocks sent is stored in the sensor.
- The sensor can verify for integrity of each data block it has sent.
- Compare the HMAC value of the data block in storage node with HMAC value of data block in sensor.
- If the values are same there is no modification in data; if values are different indicates data modification.

Verification of block d1 for eg:

If HMAC (d1) at sensor = HMAC(d1) at storage node => No data modification.

If HMAC (d1) at sensor != HMAC(d1) at storage node => Data modification.

A forged message is detected (as shown in fig 3) if a parent's calculated HMAC is inconsistent with HMAC received from the child.

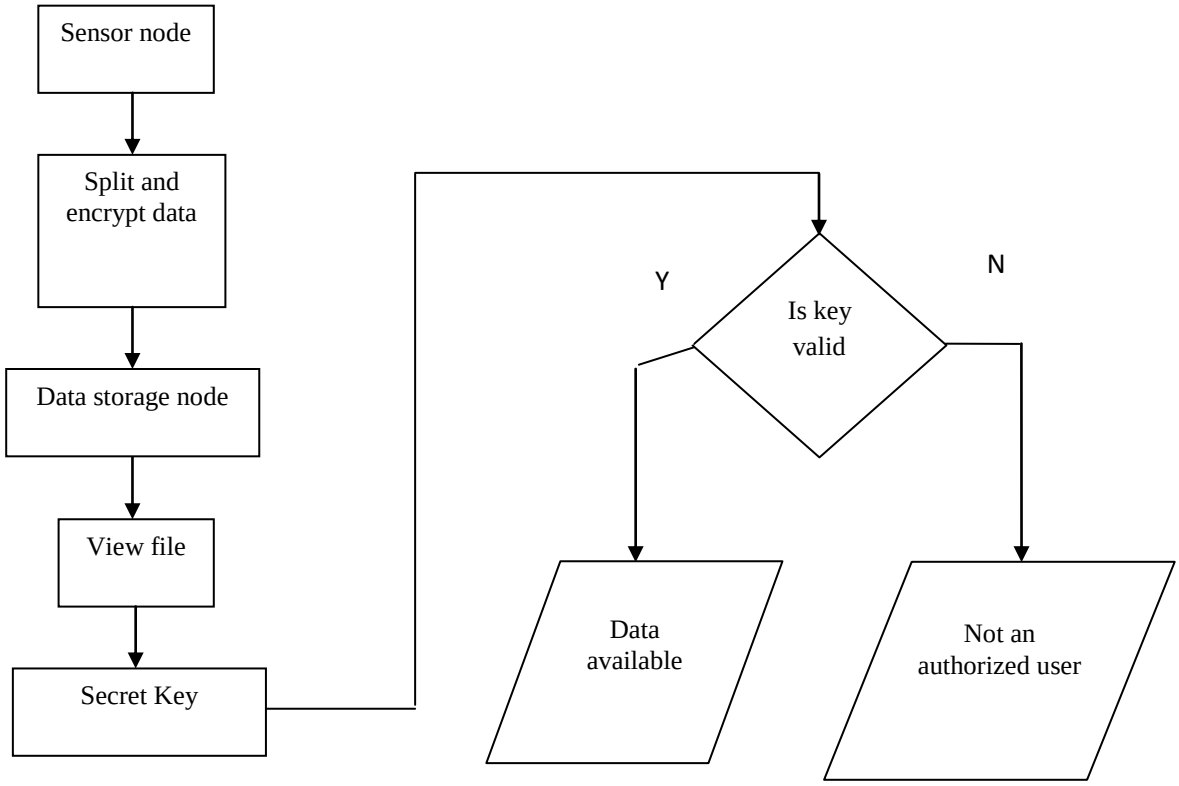

Fig 2: Flowchart to show confidentiality

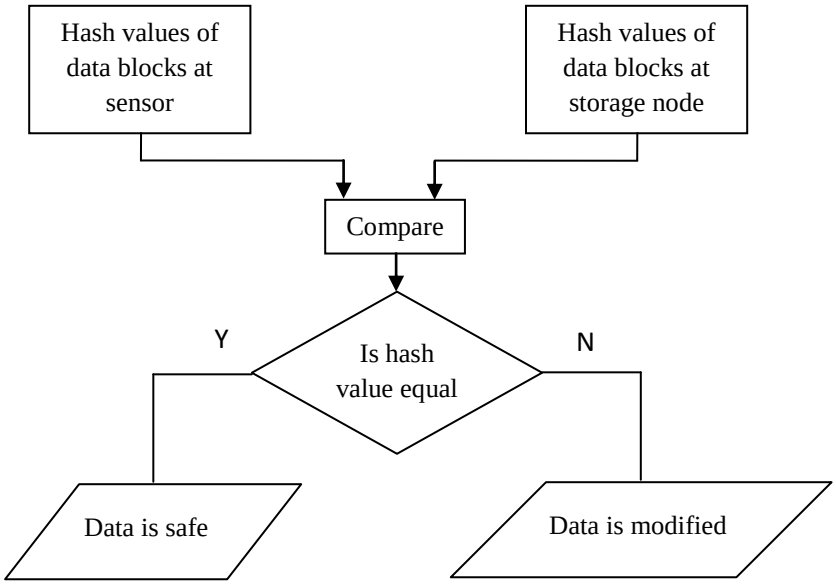

Fig 3: Flow chart to show reliability

VI. Conclusion

The benefits and problems of the intermediate tier, data storage node, in two-tiered sensor network are discussed. Confidential communication is achieved by making the storage node to provide access to authorized users only. Protection of data alteration is provided by giving sensor the capability to update the modifications. In this way, security of two-tiered sensor networks is strengthened.

References

[1] S. Ratnasamy, B. Karp, S. Shenker, D. Estrin, R. Govindan, L. Yin, and F. Yu, *"Data-centric storage in sensornets with GHT, a geographic hash table," Mobile Netw. Appl.*, vol. 8, no. 4, pp. 427–442, 2003.

[2] P. Desnoyers, D. Ganesan, H. Li, and P. Shenoy, *"Presto:A predictive storage architecture for sensor networks,"* in *roc. HotOS*, 2005, p. 23.

[3] D. Zeinalipour-Yazti, S. Lin, V. Kalogeraki, D. Gunopulos, and W. A. Najjar, "*Microhash: An efficient index structure for flash-based sensor devices,"* in *Proc. FAST*, 2005, pp. 31–44.

[4] B. Sheng, Q. Li, and W. Mao, "*Data storage placement in sensor networks,"* in *Proc. ACM MobiHoc*, 2006, pp. 344–355.

[5] B. Sheng, C. C. Tan, Q. Li, and W. Mao, *"An approximation algorithm for data storage placement in sensor networks,"* in *Proc. WASA*, 2007, pp. 71–78.

[6] B. Sheng and Q. Li, *"Verifiable privacy-preserving range query in twotiered sensor networks,"* in *Proc. IEEE INFOCOM*, 2008, pp. 46–50.

[7] J. Shi, R. Zhang, and Y. Zhang, *"Secure range queries in tiered sensor networks,"* in *Proc. IEEE INFOCOM*, 2009, pp. 945–953.

[8] B. Hore, S. Mehrotra, and G. Tsudik, *"A privacy-preserving index for range queries,"* in *Proc. VLDB*, 2004, pp. 720–731.

