Privacy and Integrity Preserving Range Queries in Sensor Networks

T. KULLAYAPPA
Dept. of Computer Science &Engineering
Chadalawada Ramanamma Engineering College

M. LAKSHMI PRASANNA KUMAR
Assistant Professor,
Dept. of Computer Science &Engineering
Chadalawada Ramanamma Engineering College

Abstract:
The architecture of two- tiered sensor networks where the storage nodes serve as an intermediate tier between sensors and a sink for storing data and processing queries has been widely adopted because of the benefits of power and storage saving for sensors as well as the efficiency of query processing. The importance of storage nodes also makes them attractive to attackers. The SafeQ a protocol that prevents attackers from gaining information from both sensor collected data and sink issue queries. SafeQ also allows a sink to detect compromise storage nodes. To preserve privacy SafeQ use a novel technique to encode both data and queries such that a storage node can correctly process encoded queries over encoded data without knowing their values. The preserve integrity have two schemes one using Merle hash trees and another using a new data structure called neighborhood chains to generate integrity verification information.

Keywords: *Encoded data, Preserving Range, SafeQ, Sensor networks*

1. Introduction

Wireless sensor networks (WSNs) have been widely deployed for various applications such as environment sensing building. We consider a two tiered sensor network architecture in which storage nodes gather data from nearby sensors and answer queries from the sink of the network. The storage nodes to serve as intermediate tier between the sensors and the sink for storing data and processing queries. Storage nodes have three main benefits to sensor networks. First sensors save the power by sending all collecting data to their closest storage node previous of sending them to the sink through long routes. Second sensors can be memory limite because data are mainly store on storage nodes. The third query processing becomes more efficient because the sink only communicates with storage nodes for queries. Several products of storage nodes such as Stargaze and RISE, are commercially available. In the inclusion of storage nodes also brings significant security challenges. The storage nodes store data receive from sensors and serve as an important role for answering queries they are more vulnerable to be compromised especially in a hostile environment. The compromiser to straggling nodes imposes significant threats to a sensor network.

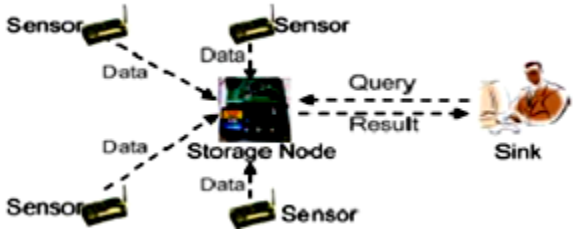


Fig. 1. Architecture of two-tired sensor networks

We want to design a protocol that prevents attackers from gaining information from both sensor collect data and sink issued queries which typically can be modele as range queries and allows the sink to detect compromised storage nodes when they misbehave. The privacy of compromising a storage nodes should not allow the attacker to obtain the sensitive information that has been and will be stored in the node as well as the queries that the storage node has receive.

2. Existing System

In the existing Wireless sensor networks once sensor nodes have been deploye there will be minimal manual intervention and monitoring. The nodes are deploye in a hostile environment and there is no manual monitoring. In the existing system the architecture of two tiere sensor networks where storage nodes serve as an intermediate tier between sensors and a sink for storing data and processing queries has been widely adopte because of the benefits of power and storage saving for sensors as well as the efficiency of query processing. The importance of storage nodes also makes them attractive to attackers.

2.1 System model

We assume that two tire sensor networks. The two tire sensor network consists of three types of nodes sensors storage nodes and a sink. Sensors are inexpensive sensing devices with limite of storage and computing power. They often massively distributes in a field for collecting physical or environmental data.

2.2 Threat Model

We assume that the sensors and the sink or trust but the storage nodes are not in a hostile environment both sensors and storage nodes can be compromise. If a sensor is compromise the subsequent collecting 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 that attacks without the use of tamper proof hardware.

2.3 Disadvantages

A compromising storage node imposes significant threats to a sensor networks.

- The attacker may obtain sensitive data that has been or store in the storage node.
- The compromising storage nodes may return forged data for a query.
- This storage node may not include all data items that satisfy the query.

3. PROPOSED SYSTEM

The proposes a scheme to preserve the privacy and integrity range of queries in sensor networks. This scheme uses the bucket partitioning for database privacy. Basic idea is to be divide the domain of data values into multiple buckets to the size of that is compute based on the distribution of data values and the location of sensor. In each time the slot of sensor collect data items from the environment places them into buckets. It has no data items in the sensor sends an encoding number which can be use by the sink to verify that the bucket is empty to a nearby storage node. Receiving the bucket ID's the storage node returns the corresponding encrypted data in all those buckets. SafeQ also allows a sink to detect compromise storage nodes. The sink is the point of contact for users of the sensor network.

4. Module Description

SafeQ

SafeQ is a protocol that prevents attackers from gaining information from both sensor collected data and sink issue queries. SafeQ also allows a sink to detect compromise storage nodes when

they misbehave. To preserve *privacy* uses a novel technique to encode both data and queries such that a storage node can correctly process encode queries over encoded data without knowing their values.

Integrity

The sink needs to detect whether a query result from a storage node includes forged data items or does not include all the data that satisfies the query. There are two key challenges to solving the privacy and integrity-preserving range query problem. A storage node needs to correctly process encode queries over encode data without knowing their actual values. A sink needs to verify that result of a query contains all the data items that satisfy the query and does not contain any forged data.

Privacy

To preserve privacy SafeQ use a novel technique to encode both data and queries such that a storage node can correctly process encoding queries over encoding data without knowing their actual values.

Range Queries

The queries from the sink are range of queries. A range query to finding all the data items collect at time slot in the range is denotes. Note that the queries in most sensor network applications can be easily modeled as range queries.

Sink

The sink is the point of contact for users of the sensor networks. 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 basing their data and return the query results to the sink.

Storage Node

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 items from the sensor nodes. The storage node cannot be view the actual value of sensor node data. If storage node trying to view the sensor node data sink detect misbehave of storage node.

Advantages

- The system is privacy- and integrity-preserving.
- SafeQ significantly strengthens the security of two-tiered sensor networks.
- In terms of effective results shown that SafeQ significantly outperforms prior art for multidimensional data in terms of both power consumption and storage space.

Privacy for One-Dimensional Data

To preserve privacy, it seems natural to have sensors encrypt data and the sink encrypt queries. However, the key challenge is how a storage node processes encrypted queries over encrypt data.

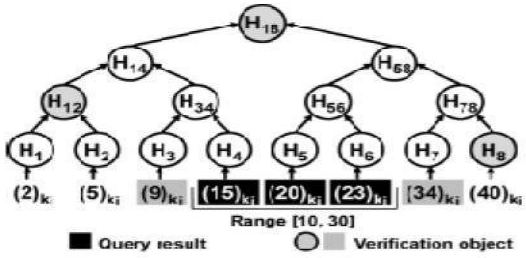


Fig. 2. Data integrity verification

4.1 Integrity for One-Dimensional Data

Our first integrity-preserving mechanism is based on Merkle hash trees. Each time a sensor sends data items to storage nodes, it constructs a Merkle hash tree for the data items. The number of data items n is not a power of 2, interim hash values that do not have a sibling value to which they may be concatenated are promote without any change up the tree until a sibling is found. If we remove the nodes H6,H7,H8 and let H58=H56=H5, the resulting unbalance tree is the Merkle hash tree for five data items.

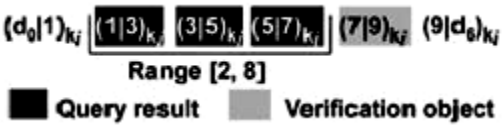


Fig. 3. Example neighborhood chain

4.2 Integrity for Multidimensional Data

To preserve the privacy of multi-dimensional data, we apply our 1-dimensional privacy preserving techniques to each dimension of multi-dimensional data. For example, sensor si collects 5 two-dimensional data items (1,11), (3,5), (6,8), (7,1) and (9,4), it will apply the 1 - dimensional privacy preserving techniques to the first dimensional values {1, 3, 6, 7, 9} and the second dimensional values {1, 4, 5, 8, 11}. Given a range query ([2,6],[3,8]), the query result QR1 for the sub-query [2,6] is the encrypted data items of (3,5),(6,8) and the query result QR2 for the sub query [3,8] is the encrypted data items of (9,4),(3,5),(6,8). Therefore the query result QR is the encrypted data items of (3,5),(6,8).

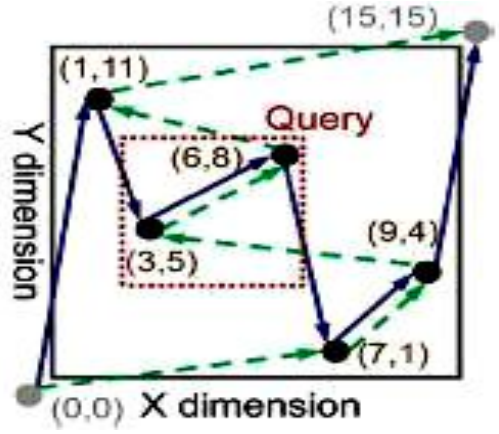
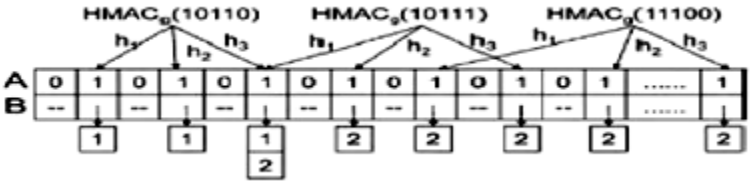


Fig. 4. Two-dimensional neighborhood chain

4.3 SafeQ Optimization

To reduce the communication cost between sensors and storage nodes, for n data items d1,....., dn, we use a Bloom filter to represent hg(p([min, d1])), hg(p([d1, d2])),.....,hg(p([dn−1, dn])), hg(p([dn,max])). Thus, a sensor only needs to send the Bloom filter instead of the hashes to a storage node. The number of bits needed to represent the Bloom filter is much smaller than that needed to represent the hashes.



5. Experiment Results

The experimental results from our side-by-side comparison show that SafeQ significantly outperforms the S&L scheme for multidimensional data in terms of power and space consumption. In the two integrity preserving schemes the neighborhood-chaining technique is better than Merkle hash tree technique in terms of both power and space consumption.

Figs. 6(d) and 7(d) show the average power consumption for a 10-min slot for a sensor and a storage nodes respectively versus the number of dimensions of the data. We observe that there are almost linear correlations between the average power consumption for both sensors and storage nodes and the number of dimensions of the data which also confirms our complexity analysis.

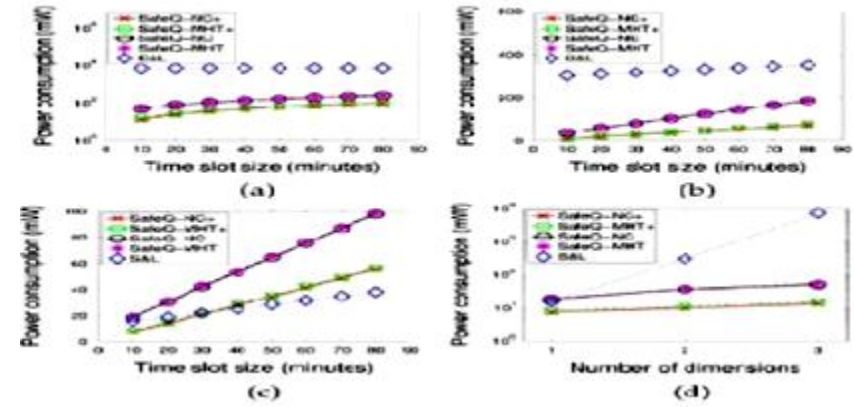


Fig. 6. Average power consumption per submission for a sensor. (a) Three-dimensional data. (b) Two-dimensional data. (c) One-dimensional data. (d) For 10 min.

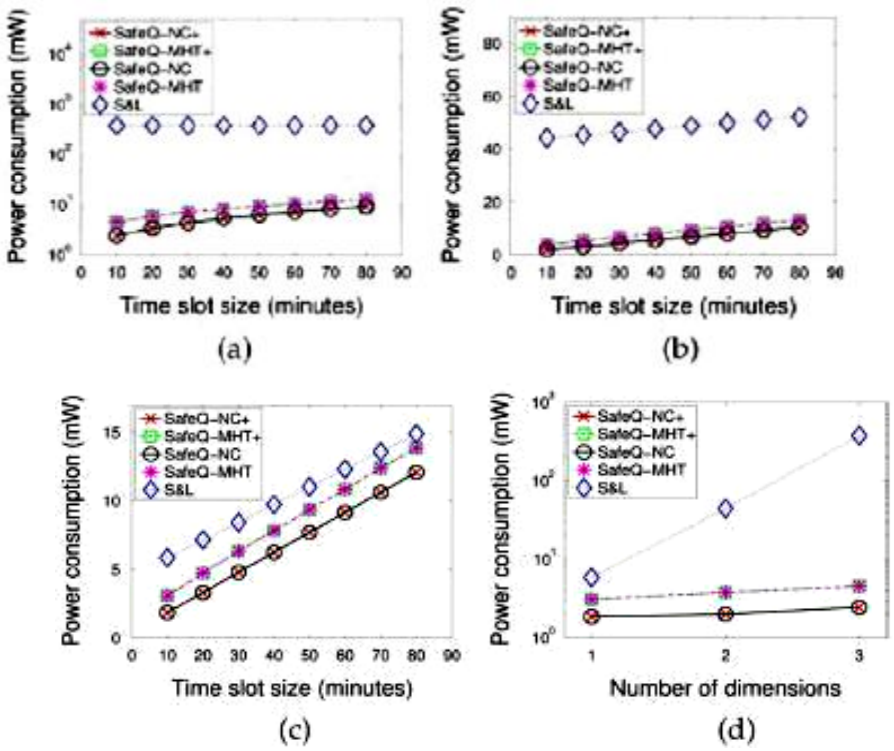


Fig. 7. Average power consumption per query response for a storage node. A. Three-dimensional data. B. Two-dimensional data. C. One-dimensional data. D. For 10 min.

Fig. 8(a) to (c) shows the average space consumption of storage nodes for three, two and one dimensions data respectively. The space consumption on storage nodes in comparison to the S &

L scheme.

6. Conclusion

We make three key contributions on this paper. We proposes SafeQ a novel and efficient protocol for handling range queries in two tiere sensor networks in a privacy and integrity preserving fashion. SafeQ uses the techniques of prefix membership verification Merkle hash trees and neighborhood chaining. In terms of security SafeQ significant strength is the security of two tiere sensor networks. In the prior art SafeQ prevents a compromising storage node from obtains reasonable estimation on the actual values of sensor collecte data items and sink issue queries.

Enhancement

In the future research on minimizing communication overhead among sensor nodes subsequently power consumption and storage space of storage nodes

References

1. Chen, F. and A. X. Liu, (2010). "SafeQ: Secure and efficient query processing in sensor networks," in Proc. IEEE INFOCOM. pp. 1–9.
2. Desnoyers, P. D., Ganesan, H. Li, and P. Shenoy,(2005)."Presto:A predictive storage architecture for sensor networks," in Proc. Hotos, 2005, p. 23.
3. Ratnasamy, S. B., Karp, S. Shenker., D. Estrin., R. Govindan., L. Yin, and F. Yu,(2003). "Data-centric storage in sensornets with GHT, a geographic hash table," Mobile Netw. Appl., vol. 8, no. 4, pp. 427–442.
4. Sheng, B. Q., Li, and W. Mao, (2006)."Data storage placement in sensor networks," in Proc. ACM MobiHoc. pp. 344–355
5. Zeinalipour-Yazti., D. S. Lin., V. Kalogeraki., D. Gunopulos, and W. A. Najjar, (2005)."Microhash: An efficient index structure for flash-based sensor devices," in Proc. FAST. pp. 31–44.