A NOVEL APPROACH FOR PRIVACY AND INTEGRITY TOP-K QUERY PROCESSING FOR TWO-TIERED SENSOR NETWORK

NEMALIDENNA PRASANNA LAKSHMI[1], M.JAGADISH[2], SK.KHALEELULLAH[3]

[1]Student, M.Tech (CSE), [2]Assistant professor, [3]Assistant professor & HOD
LINGAYAS INSTITUTE OF MANAGEMENT & TECHNOLOGY, A.P., India.

Abstract: **In this paper, focus on a two tiered sensor networks. Where resource-rich storage nodes act as an intermediate tier between sensor nodes and the sink. The storage nodes store data from their nearby sensors and process queries from the sink. In this paper, we propose SafeQ, a protocol that prevents attackers from gaining information from both sensor collected data and sink issued queries. SafeQ also allows a sink to detect compromised storage nodes when they misbehave. To preserve privacy, SafeQ uses 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. To preserve integrity, we propose two schemes — one using merkle hash trees and another using a new data structure called neighborhood chains — to generate integrity verification information so that a sink can use this information to verify whether the result of a query contains exactly the data items that satisfy the query. To improve performance, we propose an optimization technique using Bloom filters to reduce the communication cost between sensors and storage nodes.**

Keywords: Two-tiered sensors, Sensor Network, SafeQ , privacy, range queries.

INTRODUCTION

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, the sensors can be memory-limited because the data are mainly stored on storage nodes. Third, query processing becomes more efficient because the sink only communicates with storage nodes for queries. The storage node faces serious security challenge in hostile environment first, when the storage node is compromised the sensing data from the sensor nodes, the history of query requests and the corresponding query results are exposed. Second, it may cause heavy loss when the compromised sensor nodes return fake, forged or incomplete data for a query especially in military and commercial application. Therefore, developing a privacy- preserving and result-verification mechanism is of paramount importance, such that the authenticity and completeness of the query results can be verified as well as the privacy of the sensitive data is protected.

LITURATURE SURVEY

Sheng and Li ,Hacigumus et al for database privacy. 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 the 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, which can be used by the sink to verify that the bucket is empty, to a nearby storage node. 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 node. Upon receiving the bucket IDs, the storage node returns the corresponding encrypted data in all those buckets. The sink can then decrypt the encrypted buckets and verify the integrity using encoding numbers.

MODULE DESCRIPTION

SafeQ:
SafeQ is a protocol that prevents attackers from gaining information from both sensor collected data and sink issued queries. SafeQ also allows a sink to detect compromised storage nodes when they misbehave. To preserve privacy, SafeQ uses 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.

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 satisfy the query. There are two key challenges in solving the privacy and integrity-preserving range query problem. First, a storage node needs to correctly process encoded queries over encoded data without knowing their actual values. Second, a sink needs to verify that the 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 uses 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 actual values.

Range Queries
The queries from the sink are range queries. A range query ―finding all the data items collected at time-slot in the range ‖ is denoted as. 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 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. Sink can detect compromised storage nodes when they misbehave.

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 from the sensor nodes. The storage node can't view the actual value of sensor node data. If the storage node trying to view the sensor node data, sink detect misbehave of storage node.

PROPOSED SYSTEM
In the existing system, the architecture of two-tiered sensor networks, where 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. However, the importance of storage nodes also makes them attractive to attackers. In this paper, we propose SafeQ, a protocol that prevents attackers from gaining information from both sensor collected data and sink issued queries. SafeQ also allows a sink to detect compromised storage nodes when they misbehave. We propose SafeQ, a novel and efficient protocol for handling range queries in two-tiered sensor networks in a privacy- and integrity- preserving fashion.

RELATED WORK
Privacy-preserving and Secure PriSecTopk Before giving our intact result, we start with a straightforward yet ideal scheme where the storage nodes follow the designate protocol to better illustrate the privacy problem in the two-tier wireless sensor network. Then, we assume **Algorithm 1 encOPSE(D,R,m)**

1: M ← |D|; N ← |R|;
2: d←min(D)-1;r←min(R)-1; y←r+dN/2e
3: if |D|=1 then
4: cc←TapeGen(K,1lR ,(D,R,1km)); c←Rand(cc,R);return c
5: end if
6: cc←TapeGen(K,1lR ,(D,R,0ky)); x←HGD(D,R,y;cc)
7: if m≤ x then
8: D←{d+1,. . .,x}; R←{r+1,. . .,y}
9: else
10: D←{x+1,. . .,d+M}; R←{y+1,. . .,r+N}
11: end if
12: return encOPSE(D,R,m)

a more general attack model that the compromised storage nodes do not follow the designate protocol but return the fake/incomplete query result in addition to learn additional information over the sensing data and query request. A. PriSecTopk I: Basic Scheme With respect to the plaintext attack model, PriSecTopk I enables the privacy of sensing data and query request by encrypting the data items with revised order-preserving and traditional symmetric encryption respectively. 1) PriSecTopk I: To provide a privacy guarantee against the attack on the plaintext of sensing data and query request as well as the efficiency requirement, we utilize revised order-preserving encryption scheme and traditional symmetric encryption to encrypt the sensing data and query request. The order-preserving symmetric encryption scheme (OPS E) is based on the observation that any order-preserving function g from {1, . . . , M} to {1, . . . , N} can be uniquely represented by a combination of M out of N order items. The whole scheme to achieve top-k query over encrypted data in two-tier wireless sensor network is as follows. Before deployed, each sensor node u shares with the user a unique secret key ku which we call the node's individual key. Also the user and the sensor nodes have the same pseudo-random function f(), corresponding seed and collision resistant hash function π. The sensor initiates the scheme by generating random keys x, y from the pseudo-random function f(). After gaining the sensing data du at the time-slot t, the sensor node u calculates the score su for the sensing data. Let E be a semantically secure symmetric encryption algorithm. Each sensor node u encrypts the score su by orderpreserving symmetric encryption as shown in Algorithm 1 to gain the ciphertext of the score OPS E(su) and then encrypt node ID using key fx(t), compute the keyed-hash for time stamp t. Sensor node u sends the encrypted data along with Message Authentication Code to the closest storage node S . u → S : Efx(t)OPS E(su), Efy(t)(idu), πx(y), MAC(ku, m) where m = Efx(t)OPS E(su)|Efy(t)(idu)|πx(y) To encrypt the query request Qq =< q, t > where q

is the query code and t is the time stamp to query. With the pseudo-random function f() and the corresponding seed, user generates the secret key x to encrypt the query time stamp with the one-way hash function π and then sends the secret key fx(t) along with the ciphertext of the query time stamp to the storage node. user → S : q, πx(t), fx(t) With the ciphertext of the query time stamp, storage node locates the matching list of the sensing data via πx(t), uses fx(t) to achieve the OPS E sensing scores and then sorts the OPS E sensing scores to fetch the top k sensing data records. Storage nodes send back the top-k encrypted sensing records i.e. Efx(t)OPS E(su),Efy(t)(idu),πx(y),MAC(ku, m) to the user. Upon receiving the query response from the storage node, the user first utilizes the pseudo-random function and the corresponding seed shared with sensor node to generate the secret key y and gain fy(t) to decrypt sensor node ID, then locates the key for the each sensor node with which it can verify the MACs returned along with every sensing data record. After decrypting the OPS E(su) by order-preserving decryption scheme, the user finally obtains the top-k sensing scores as the query result. 2) Analysis: We analyze the basic scheme from two aspects. Privacy: As for query privacy, traditional symmetric key encryption techniques could be properly utilized by the user when issuing query request which are not within the scope of this paper. The data privacy is well protected not only by the symmetric key encryption but also the order-preserving encryption scheme to guarantee the top-k query. The essence of order-preserving encryption can be seen as a combination of the size of domain M out of the size of range N ordered items. An adversary can only break the encryption scheme by performing a brute force search over all the possible combination of M out of N. If the security level is chosen to be 80 bits, then it is suggested to choose M= N/2>80 so that the total number of combination will be greater than 2 80. Therefore, the data privacy and query privacy are well protected in our basic scheme, while it is an unsolved privacy leakage problem of result privacy. Efficiency: With the help of order-preserving encryption scheme, the storage node can process the top-k query as efficiently as for the unencrypted sensing scores. In the encryption and decryption process, the number of recursive calls is at most log N + 1 in the worst-case and at most 5 log M + 12 on average. As for memory addition, PriSecTopk I do not ask for additional memory except for the MACs.

CONCLUSION

In this paper, proposed the problem of top-k query on time slot data set in two-tier wireless sensor network, and establish a set of privacy and correctness requirements for such a secure top-k scheme to become practical. We propose Two-tiered SafeQ schemes meeting different privacy and correctness requirements in consideration of two levels of threat models. Thorough analysis investigating privacy, detection rate and efficiency guarantee of proposed scheme is given, and experiments on the real-world dataset further show the efficiency of proposed schemes.

REFERENCES

[1] Fei Chen and Alex X. Liu,‖Privacy- and Integrity-Preserving Range Queries in Sensor Networks‖,**IEEE/ACM TRANSACTIONS ON NETWORKING, 2012.**
[2] W. Zhang, H. Song, S. Zhu, and G. Cao. Least privilege and privilege deprivation: towards tolerating mobile sink compromises in wireless sensor networks. in MobiHoc05, pp. 378C389. ACM,2005.
[3] M. Shao, S. Zhu, W. Zhang, and G. Cao.pDCS: Security and privacy support for data-centric sensor networks. in INFOCOM07, pp. 1298C1306. IEEE, 2007
[4] S. Ratnasamy, B. Karp, S. Shenker. Data-centric storage in sensornets with GHT, a geographic hash table. Mobile Networks and Applications, 2003,8(4):427-442
[5] Willow Technologies: SPB400-STARGATE GATEWAY. http://www.willow.co.uk/html/spb400- stargate gateway.html
[6] Sheng B, Li Q. Verifiable privacy-preserving range query in two tiered sensor networks. In INFOCOM'08, pp.46-50. IEEE 2008
[7] Shi J, Zhang R, Zhang Y. Secure range queries in tiered sensor networks. In INFOCOM'09. pp.197-206. IEEE, 2009
[8] Fei C, Alex L. SafeQ: Secure and Efficient query processing in sensor networks. In NFOCOM'10. IEEE, 2010
[9] Rui Z, Jing S, Yunzhong L, et al. Verifiable fine-grained top-k queries in tiered sensor networks. In INFOCOM'10. IEEE, 2010
[10] Samuel Madden. Intel Lab Data: sensor readings. http://db.csail.mit.edu/labdata/labdata.html
[11] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu. Order-preserving encryption for numeric data. In S IGMODE004, pp.563-574. ACM,2004
[12] A. Boldyreva, N. Chenette, Y. Lee and A. O'Neill. Order-preserving symmetric encryption. In Eurocrypt009, PP.224-241. Springer,2009
[13] A. Boldyreva, N. Chenette, A. O'Neill. Order-preserving Encryption Revisited: Improved Security Analysis and Alternative Solutions. In ACC011, PP.578-595.Springer,2011
[14] H. Hacigümüş, B. Iyer, C. Li, and S. Mehrotra, "Executing SQL over encrypted data in the database-serviceprovider model," in Proc. ACM SIGMOD, 2002, pp. 216–227.
[15] R. Merkle, "Protocols for public key cryptosystems," in Proc. IEEE S&P, 1980.