Journal of Engineering Sciences
Vol 10, Issue 9, Sept /2019
ISSN NO:0377-9254

PRIVACY AND INTEGRITY PRESERVING TOP-K QUERY PROCESSING FOR TWO-TIERED SENSOR NETWORK

[1]K. SRAVYA , [2]P. MURALI PONAGANTI
[1]MCA Student, [2]Associate Professor
DEPARTMENT OF MCA
SREE CHAITANYA COLLEGE OF ENGINEERING, KARIMNAGAR

ABSTRACT: Privacy and integrity have been the main road block to the applications of two-tiered sensor networks. The storage nodes, which act as a middle tier between the sensors and the sink, could be compromised and allow attackers to learn sensitive data and manipulate query results. Prior schemes on secure query processing are weak, because they reveal non-negligible information, and therefore, attackers can statistically estimate the data values using domain knowledge and the history of query results. In this paper, we propose the first top-k query processing scheme that protects the privacy of sensor data and the integrity of query results. To preserve privacy, we build an index for each sensor collected data item using pseudo-random hash function and Bloom filters and transform top-k queries into toprange queries. To preserve integrity, we propose a data partition algorithm to partition each data item into an interval and attach the partition information with the data. The attached information ensures that the sink can verify the integrity of query results. We formally prove that our scheme is secure under IND-CKA security model. Our experimental results on real-life data show that our approach is accurate and practical for large network sizes.

I. INTRODUCTION

With the popularity of laptops, cell phones, PDAs, GPS devices, RFID, and intelligent electronics in the post-PC era, computing devices have become cheaper, more mobile, more distributed, and more pervasive in daily life. It is now possible to construct, from commercial off the shelf (COTS) components, a wallet size embedded system with the equivalent capability of a 90's PC. Such embedded systems can be supported with scaled down Windows or Linux operating systems. From this perspective, the emergence of wireless sensor networks (WSNs) is essentially the latest trend of Moore's Law toward the miniaturization and ubiquity of computing devices.

Typically, a wireless sensor node (or simply sensor node) consists of sensing, computing, communication, actuation, and power components. These components are integrated on a single or multiple boards, and packaged in a few cubic inches. With state-of-the-art, low-power circuit and networking technologies, a sensor node powered by 2 AA batteries can last for up to three years with a 1% low duty cycle working mode. A WSN usually consists of tens to thousands of such nodes that communicate through wireless channels for information sharing and cooperative processing. WSNs can be deployed on a global scale for

environmental monitoring and habitat study, over a battle field for military surveillance and reconnaissance, in emergent environments for search and rescue, in factories for condition based maintenance, in buildings for infrastructure health monitoring, in homes to realize smart homes, or even in bodies for patient monitoring.

After the initial deployment (typically ad hoc), sensor nodes are responsible for self-organizing an appropriate network infrastructure, often with multi-hop connections between sensor nodes. The onboard sensors then start collecting acoustic, seismic, infrared or magnetic information about the environment, using either continuous or event driven working modes. Location and positioning information can also be obtained through the global positioning system (GPS) or local positioning algorithms. This information can be gathered from across the network and appropriately processed to construct a global view of the monitoring phenomena or objects. The basic philosophy behind WSNs is that, while the capability of each individual sensor node is limited, the aggregate power of the entire network is sufficient for the required mission.

In a typical scenario, users can retrieve information of interest from a WSN by injecting queries and gathering results from the so-called base stations (or sink nodes), which behave as an interface between users and the network. In this way, WSNs can be considered as a distributed database. It is also envisioned that sensor networks will ultimately be connected to the Internet, through which global information sharing becomes feasible

II. LITERATURE SURVEY

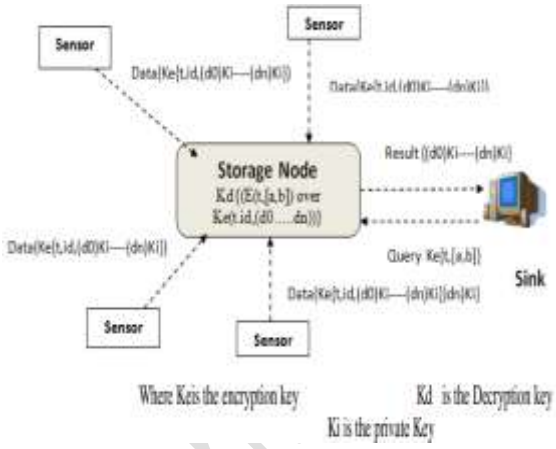
A Privacy- and integrity-preserving range query in WSNs has been studied in the recently [4], [5]. Sheng and Li proposed a scheme to preserve the privacy and integrity of range queries in sensor networks [4]. Proposes SafeQ [6], a protocol that prevents attackers from gaining information from both sensor 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 value. To preserve integrity, we propose a Merkle hash tree [9] and new data structure called neighborhood chains that allow a sink to verify whether the result of a query contains exactly the data items that satisfy the query. In addition, it proposes a solution to adapt SafeQ for event-driven sensor networks.

Sheng and Li scheme uses the bucket-partitioning idea proposed by Hacigumus et al [8] 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.

III. PROPOSED MODEL

The proposed work uses two techniques for privacy and integrity preserving entire range queries in sensor network. During privacy preserving sensor send data in the form of 3-tuple {id, t, (d0,……., dn)} and sink also send data in the form of 2-tuple {t, [a.b]} to storage node, all this data items are encrypted using paillier cryptosystem. During integrity preserving wecan use a Merkle hash tree and neighborhood chain for verifying information given by storage node to sink. The proposed diagram of two-tiered architecture of wireless sensor network as shown in Fig



Two Tiered Architecture of WSN.

IV. IMPLEMENTATION

- **Source**

  In this module, the Source activates all the sensors and assigns temperatures to all the cells, uploads their data to the particular storage node. It will store in cell. The service provider, can view the attacked file in the storage node, He can replace the injected malicious files to its Original file in the cells, and can inform to the authority about the malicious files in the cells.

- **Router**

  The Authority can issue queries to retrieve the sensor readings. The middle tier is composed of a small number of storage-abundant nodes, called *storage nodes*. The bottom tier consists of a large number of resource-constrained ordinary sensors that sense the environment. The authority issues proper queries to retrieve the desired portion of sensed data. We restrict ourselves in this system to discussing top-*k* query, which is one of the most intuitive and commonly used queries.

- **Storage Node**

  The storage node keeps a copy of received sensor readings and is responsible for answering the queries from the authority. The temperature will be stored with their tags such as node name, temperature, status, digital sign, and modified temperature in Storage Node, Also the data file will be also stored with their tags such as node name, file name, secret key, status, digital sign, and with ranks in Storage Node, Storage nodes are storage-abundant, can communicate with the authority *via* direct or multi-hop communications, and are assumed to know their affiliated cells. The Storage Node can also view the Attacker details.

- **Cell**

In This module the sensor nodes are usually partitioned into disjoint groups, each of which is associated with a storage node. Each group of sensor nodes is called a *cell*. The sensor nodes in a cell form a multi-hop network and always forward the sensor readings and file details to the associated storage node.

End User

In this module, the End user can access the top k file details and top k temperatures of the cells in associated Storage node(SN1,SN2,SN3) , The End user can request and gets file contents response from the corresponding Storage node. If the file name and secret key is correct then the enduser is getting the file response from the Authority and storage node.

- **Topk Query Processing**

The Topk queries in a centralized uncertain database, which provides a good background for the targeted distributed processing problem. The query answer can be obtained by examining the tuples in descending ranking order from the sorted table (which is still denoted as T for simplicity). We can easily determine that the highest ranked k tuples are definitely in the answer set as long as their confidences are greater than p since their qualifications as PT-Topk answers are not dependent on the existence of any other tuples.

- **Attacker**

Attacker is one who is injecting the file by adding malicious data to the corresponding storage node. The Attacker can also modify the temperature in the cells.

V. CONCLUSIONS

In this paper, we propose the first secure top-k query processing scheme that is secure under the IND-CKA security model. The data privacy is guaranteed by encryption as well as a careful generation of data indexes. We make two key contributions in this paper. The first contribution is to transform a top-k query to a top-range query and adopt membership testing to test whether a data item should be included in the query result or not. This transformation allows the storage node to find k smallest or biggest data values without using numerical comparison operations, which is a key technique for the scheme to be secure under the INDCKA security model. The second contribution is the data partition, index selection, and interval information embedding technique. This technique guarantees that at least one data item of each sensor collected data will be included in a query result and allows the sink to verify the integrity of query result without extra verification objects. Experiments show that the proposed scheme is bandwidth efficient and highly practical. The techniques proposed in this paper can be potentially useful for many other applications as well.

REFERENCES

[1] P. Desnoyers, D. Ganesan, H. Li, M. Li, and P. Shenoy, "Presto: A predictive storage architecture for sensor networks," in Proc. 10th HotOS, 2005, pp. 12–15.

[2] S. Ratnasamy et al., "Data-centric storage in sensornets with GHT, a geographic hash table,"

Mobile Netw. Appl., vol. 8, no. 4, pp. 427–442, Aug. 2003.

[3] B. Sheng and Q. Li, "Verifiable privacy-preserving range query in twotiered sensor networks," in Proc. 27th INFOCOM, Apr. 2008, pp. 46–50.

[4] B. Sheng, Q. Li, and W. Mao, "Data storage placement in sensor networks," in Proc. 7th ACM MobiHoc, May 2006, pp. 344–355.

[5] 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. 4th USENIX FAST, Dec. 2005, pp. 31–44.

[6] Stargate Gateway (SPB400), accessed on 2011. [Online]. Available: http://www.xbow.com.

[7] Rise Project. accessed on 2011. [Online]. Available: http://www.cs.ucr.edu/~rise.

[8] I. F. Ilyas, G. Beskales, and M. A. Soliman, "A survey of top-k queries processing techniques in relational database system," ACM Comput. Surv., vol. 40, no. 4, pp. 11:1–11:58, Oct. 2008.

[9] A. S. Silberstein, R. Braynard, C. Ellis, K. Munagala, and J. Yang, "A sampling-based approach to optimizing top-k queries in sensor networks," in Proc. 22nd ICDE, Apr. 2006, p. 68.

[10] R. Zhang, J. Shi, Y. Zhang, and X. Huang, "Secure top-k query processing in unattended tiered sensor networks," IEEE Trans. Veh. Technol., vol. 63, no. 9, pp. 4681–4693, Nov. 2014.

[11] E.-J. Goh, "Secure indexes," Cryptol. ePrint Archive, Rep. 2003/216. [Online]. Available: http://eprint.iacr.org/2003/216/

[12] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 1, pp. 222–233, Jan. 2014.