"A Survey On An Epoch Based Secure Data Aggregation And Authentication Scheme For Range Query Result Evaluation"

Guruprasad
M Tech Student
Dept of CSE
VTU, East west Institute of Technology
Bangalore, India

Prasanna G
M Tech Student
Dept of CSE
VTU, EWIT
Bangalore, India

Dr.Arun Biradar
HOD
Dept of CSE
VTU, EWIT
Bangalore,India

***Abstract*— We consider an unattended tiered sensor network (UTSN) which consisting of two tiers. An upper tier is associated with a resource rich master nodes and the lower tier consisting of resource poor sensor nodes. Whenever there is a query from the network owner the master nodes will answer for the query on behalf of the sensor nodes. The sensor nodes will just perform the sensing operation and send the data to the master nodes, the master nodes store that sensed data to the network owner when needed through ad-hoc communication. Such a co-operative data storage and query processing offers a number of advantages over traditional network. Storage only on the master nodes raises the security concerns, data confidentiality and result correctness in case of hostile environment. Incomplete query results may also occur due to leakage of data from the master nodes here we take an example of the multidimensional range queries to investigate the secure co-operative data storage and query processing in UTSN. So we introduce some suite of epoch schemes and we are generating an authentication from network owner to master nodes for data confidentiality and completeness of the query.**

Keywords-UTSN,Epoch,Range queries,Data storage and query processing.

I. INTRODUCTION

Unattended Tiered Sensor Networks is a form of wireless sensor networks it works with an intermittent sink(starting And stopping at particular intervals)presence. An UTSN consists of two tiers an upper tier is associated with a resource rich nodes which are called as master nodes and the lower tier consisting of resource poor nodes called as sensor nodes. The sensor nodes will perform the sensing task and send the data to the near by master nodes, whenever the data is needed by the network owner it will query the master nodes for the data the master nodes will answer those queries on behalf of the sensor nodes via an ad-hoc communication connection(e.g. Military applications). So in this kind of situation there is necessity in the network for the data storage such that the data continuously produced by the sensor network must be stored in the network. The UTSN are expected to be deployed in the remote and extreme environments such as volcano's, oceans and battlefields. In this paper we take an multidimensional range queries as an example to investigate the unattended tired sensor network range queries are nothing but the important queries which ask for the data with one or more attributes falling under some specified ranges for example the vehicles which are moving in the range 60-120Kmps.

In traditional sensor network , such range queries are supported using pre-computed indices. Indices trade-off some initial pre-computation cost to achieve a significantly more efficient querying capability. For sensor networks, we assert that a centralized index for multi-dimensional range queries may not be feasible for energy-efficiency reasons (as well as the fact that the access bandwidth to this central index will be limited, particularly for queries emanating from within the network). Rather, we believe, there will be situations when it is more appropriate to build an in network distributed data structure for efficiently answering multi-dimensional range queries.

There are two ways for the network owner to access the data generated by sensor nodes in the two-tier WSN. First, sensor nodes send their data to the nearby master nodes which in turn forward the data immediately along an upper-tier multi-hop path to the network owner. This approach requires a stable always on communication connection from some master node(s) to the network owner, which is impossible or prohibitive to maintain in extreme and hazardous environments such as oceans, animal habitats, and battlefields. This approach may also consume unnecessary energy of master nodes if the network owner is only interested in a small portion of the potentially huge amount of data produced over time. The second approach takes advantage of the rapid progress in storage technology and views the WSN as a storage system . In particular, it is now possible to equip each of the relatively fewer master nodes with several gigabytes of NAND flash storage for a few tens of dollars, though it may remain economically prohibitive to furnish each sensor node with large flash storage. Master nodes then collect data from nearby sensor nodes and store them locally for extended periods of time. The network owner can query data through an on-demand communication link (e.g., a satellite link) to some master node(s).This paper is concerned with the second approach. A WSN may need to support many types of data

IJERT

queries. In this paper, we focus on supporting range queries which are an important and common type of queries in WSNs and ask for data within a certain range The reliance on master nodes for data storage and query processing raises serious security concern. In particular, many target application environments of WSNs such as military and homeland security scenarios are unattended and hostile in nature. Master nodes are attractive targets of attack and might be compromised by the adversary. Compromised master nodes will leak sensitive data such as intrusion events to the adversary. A sound scheme is thus required to let master nodes store encrypted data for which they do not hold the decryption keys, while enabling efficient query processing at the same time. In addition, the adversary may instruct compromised master nodes to return juggled and/or incomplete data in response to range queries from the network owner. Such attacks are obviously more subtle and harmful than blind DoS attacks on the WSN, especially when the query results are used as the basis for making critical decisions such as military actions. The network owner thus cannot accept the query results at their face value. Instead, it should be able to verify that all the data in the result originated from the purported sources and have not been tampered with, and query result is both *authentic* and *complete*. The term *authentic* means *complete* means that the result includes all the data satisfying the query.

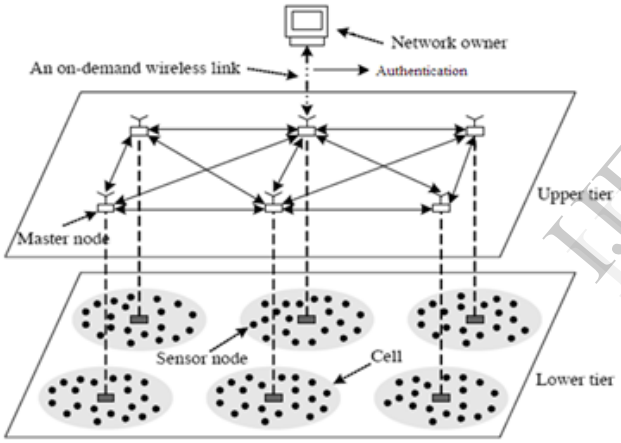

Figure 1. An UTSN with Data Storage and Query Processing

The figure.1.shows an unattended tired sensor netwoks with the data storage at the master nodes and the query processing. And the woking will be perfomed by the nodes as we said above, Aiming at one-dimensional range queries, these schemes [7], [11], [12] could ensure data confidentiality and also enable query-result authenticity and completeness verification with different communication and computation overhead. The only piece of work on secure multidimensional queries [13] relies on a common key shared among all sensor nodes, which is unfortunately vulnerable to compromised sensor nodes: the adversary can recover the common key after compromising any sensor node, whereby to further recover the original data. A more sound solution to secure multidimensional range queries thus remains an open challenge.

2. RELATED WORK

Data storage models of sensor networks have been widely discussed in prior research. Early work on this topic considers the extreme cases, archiving all data on the sink or each sensor locally. New data storage system is designed by introducing an intermediate tier between the sink and sensors, that can cache data, process query, and provide a more efficient access to the data collected by sensor networks. This paper considers the same system model, where some storage nodes are deployed as the intermediate tier and responsible for data archival and query response. In fact, this kind of special nodes have been manufactured, e.g., Star Gate and RISE. In Mathur et al. also attached external flash memory to sensor nodes and give a comprehensive evaluation of the performance. In addition, MicroHash is a file system specifically designed for sensor nodes with large storage size. In our previous work, we proposed an optimal deployment strategy of storage nodes in order to maximize performance improvement.

Here we investigates techniques to secure multidimensional range queries in UTSNs against possibly compromised master nodes. We employ the bucketing technique [14], [15] to achieve data confidentiality and also query-result authenticity verification while ensuring efficient query processing. Our major contributions are a suite of novel techniques for the network owner to verify query-result completeness. In particular, our first construction is a deterministic approach that is an extension of the technique in [7] to multidimensional cases and also serves as our benchmark. It allows the network owner to immediately catch misbehaving master nodes at the cost of high communication overhead growing exponentially with the number of dimensions (or queriable data attributes). We then present two novel probabilistic techniques with significantly less communication overhead, including a spatial crosscheck technique and a temporal crosscheck technique. The former aims to create some relationships among data generated by sensor nodes affiliated with the same master node, while the latter aims to embed some relationships among data produced in different time periods. These two techniques can collectively allow the network owner to verify with overwhelming probability whether a query result is complete by examining the spatial and temporal relationships among the returned data. We further propose a random-probing technique as a complement to spatial and temporal crosscheck techniques to cope with compromised sensor nodes. With our countermeasures in place, compromised master nodes have to return authentic and complete query results to avoid being detected. Built upon symmetric cryptographic primitives, our techniques are shown to be very effective and efficient through comprehensive theoretical analysis and performance evaluations.

IJERT

2.1 Network Model

We consider a large scale two-tier sensor network with thousand of resource poor sensor nodes and relatively some resource rich master nodes, as shown in Fig. 1. Master nodes have abundant resources in storage, energy (a solar panel and/or heavy-duty rechargeable batteries), and computation; they also communicate in a multi-hop fashion via relatively long-range high-rate radios. In contrast, sensor nodes are more constrained in storage, energy, computation, and communication capabilities. Each master node is in charge of a physical region of the network field, called a *cell*. Sensor nodes in a cell are affiliated with the master node in that cell. Here we follow the conventional assumption that master nodes and sensor nodes know their respective geographic locations and also which cell they are in, which can be realized by many existing techniques such as [12, 13]. Depending on concrete applications, the cells of two neighboring master nodes may overlap, in which case sensor nodes in the overlapping region are affiliated with both master nodes. We do not assume an always-on communication connection to the external network owner. Instead, the network owner can query data by an on demand wireless link (e.g., a satellite link) connected to some master node(s). To prevent storage overflow of master nodes, mobile sinks [14] can also be periodically (e.g., quarterly) dispatched to collect data and empty the storage of master nodes.

Here we assume that time is divided into *epoches* and that sensor and master nodes are loosely synchronized. At the end of each epoch, each sensor node submits to its master node all the data (if any) it produced during that epoch. Without loss of generality, we subsequently focus on a cell C with N sensor nodes $\{S_i\}_i^N=1$ and a compromised (yet undetected) master node M. It is worth noting that all the operations also apply to all the other cells with or without compromised master nodes.

2.2 Multidimensional Range queries

Event data generated by sensor nodes can generally be described as a tuple of attribute values $\{Aj\}_j^d=1$, where $d\geq 1$ depends on concrete sensor network applications. Each attribute Aj represents a sensor reading or an aspect of the event such as the weight of an observed object, its location, its speed and moving trajectory, or its appearance or lingering time. Let $A\subseteq \{Aj\}_j^d=1$ be a subset of attributes the network owner is interested in. For sake of simplicity, we will focus on the following type of primitive multidimensional range queries,

$$(\text{cell} = C) \wedge (\text{epoch} = t) \wedge_{Aa\epsilon A} (l_a \leq A_a \leq h_a); \quad (1)$$

where C and t denote the cell ID and the interested epoch, respectively, and [la; ha] is the interested range of attribute Aa. For other types of range queries which, for example, involve multiple epoches and/or cells or the union of attributes, they can be converted into multiple primitive range queries. Our work can also be easily extended to support range queries concerning specific sensor nodes.. Note that work aims at single-attribute (or one dimensional) range queries range queries, which are a special case of ours with only one queriable atttibute(i.e., d=1).

2.3 Adversary Model

Tremendous efforts have been made to secure sensor network communications, see for example [15–22]. This paper focuses on secure multidimensional range queries, an open challenge. We resort to the existing rich literature for other important issues such as key management, secure routing, broadcast authentication, secure localization, DoS mitigation, and particularly secure and reliable message transmissions.

Although the adversary may directly compromise sensor nodes to read their data and manipulate their behavior, it is much more tempting to take over master nodes for their significant roles in the two-tier sensor network. The adversary is assumed to have compromised some master nodes whereby to launch attacks against data confidentiality and query-result authenticity and completeness. The adversary may also compromise sensor nodes to aid compromised master nodes. We, however, follow the conventional assumption that non compromised sensor nodes are always the majority. Since every master node is only responsible for its own cell, the collusion of compromised master nodes will not do more harm. Our subsequent discussion thus concentrates on one compromised master node M in charge of a cell with N sensor nodes $\{S_i\}_i^N=1$.

In the confidentiality preserving rage queries we need to illustrate how to realize data confidentiality, query result authentication and efficient query evaluation. To ensure data confidentiality against *M.* it is necessary to store the encrypted data at *M* for which *M* has no decryption keys. In the existing cases they use the bucketing technique to strike the balance between data confidentiality and query efficiency.

In the query result completeness section verification technique we present a set of schemes for a network owner to verify the completeness of the query the word complete means that the result includes all the data which satisfying the query. Without loss of generality. we still consider cell C with sensor nodes $\{S_i\}_i^N=1$.and a compromised master node *M.* For clarity only, we first temporarily ignore compromised sensor nodes and then discus their impact and corresponding defenses in the impact of compromised sensor nodes and random probing.

The impact of compromised nodes contains the two steps such as Disobey the verification operation and Return the data from the compromised sensor nodes. We further propose a random probing scheme as a complement to spatial and temporal crosscheck [12] schemes, in which the network owner probes some random chosen nodes of which no data were returned in the query result.

IJERT

CONCLUSION

So in this paper we have seen the survey of the work related to reduction in the delay, authentication and the compklteness of the query, efficient query result evaluation with low cost and the network owner authentication in the unattended tiered sensor networks with the help of the multidimensional range queries and in future we will define specific efficient enhancing steps for this scheme.

REFERENCES

[1] R. Zhang, J. Shi, and Y. Zhang, "Secure multidimensional range queries in sensor networks," in *ACM MobiHoc'09*, New Orleans, LA, May 2009.

[2] R. D. Pietro, L. V. Mancini, C. Soriente, A. Spognardi, and G. Tsudik, "Catch me (if you can): Data survival in unattended sensor networks," in *PerCom'08*, Hong Kong, China, Mar. 2008, pp. 185–194.

[3] D. Ma, C. Soriente, and G. Tsudik, "New adversary and new threats: security in unattended sensor networks," *IEEE Network*, vol. 23, no. 2, pp. 43-48, 2009.

[4] P. Desnoyers, D. Ganesan, and P. Shenoy, "TSAR: A two tier sensor storage architecture using interval skip graphs," in *ACM SenSys'05*, San Diego, California, USA, Nov. 2005, pp. 39–50.

[5] B. Sheng, Q. Li, and W. Mao, "Data storage placement in sensor networks," in *ACM MobiHoc'06*, Florence, Italy, May 2006. [6] M. Shao, S. Zhu, W. Zhang, and G. Cao, "pDCS: Security and privacy support for data-centric sensor networks," in *IEEE INFOCOM'07*, Anchorage, Alaska, USA, May 2007, pp. 1298–1306.

[7] B. Sheng and Q. Li, "Verifiable privacy-preserving range query in sensor networks," in *IEEE INFOCOM'08*, Phoenix, AZ, Apr. 2008, pp. 46–50.

[8] O. Gnawali, K.-Y. Jang, J. Paek, M. Vieira, R. Govindan, B. Greenstein, A. Joki, D. Estrin, and E. Kohler, "The tenet architecture for tiered sensor networks," in *ACM SenSys'06*, Boulder, Colorado, USA, Oct. 2006.

[9] X. Li, Y. J. Kim, R. Govindan, and W. Hong, "Multi-dimensional range queries in sensor networks," in *ACM SenSys'03*, Los Angeles, California, USA, Nov. 2003, pp. 63–75.

[10] S. R. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TinyDB: an acquisitional query processing system for sensor networks," *ACM Trans. Database Syst.*, vol. 30, no. 1, pp. 122–173, Mar. 2005.

[11] J. Shi, R. Zhang, and Y. Zhang, "Secure range queries in tiered sensor networks," in *IEEE INFOCOM'09*, Rio de Janeiro, Brazil, Apr. 2009.

[12] , "A spatiotemporal approach for secure range queries in tiered sensor networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 264–273, Jan. 2011.

[13] F. Chen and A. Liu, "SafeQ: Secure and efficient query processing in sensor networks," in *INFOCOM'10*, San Diego, CA, Mar. 2010, pp. 1–9.

[14] H. Hacig¨um¨us¸, B. Iyer, C. Li, and S. Mehrotra, "Executing SQL over encrypted data in the database-service-provider model," in *ACM SIGMOD'02*, Madison, Wisconsin, 6 2002, pp. 216–227.

[15] B. Hore, S. Mehrotra, and G. Tsudik, "A privacy-preserving index for range queries," in *VLDB'04*, Toronto, Canada, Aug. 2004, pp. 720–731.

[16] D. Liu, P. Ning, and W. Du, "Attack-resistant location estimation in
sensor networks," in *IPSN'05*, Los Angeles, CA, Apr. 2005.

[17] Y. Zhang, W. Liu, Y. Fang, and D. Wu, "Secure localization and
authentication in ultra-wideband sensor networks," *IEEE J. Sel. Areas Commun., Special Issue on UWB Wireless Communications – Theory and Applications*, vol. 24, no. 4, pp. 829–835, Apr. 2006.

[18] H. Chan, A. Perrig, and D. Song, "Random key predistribution schemes for sensor networks," in *IEEE S&P'03*, Oakland, CA, May 2003.

[19] S. Zhu, S. Setia, and S. Jajodia, "LEAP: Efficient security mechanisms for large-scale distributed sensor networks," in *ACM CCS'03*, Washington, DC, Oct. 2003, pp. 62–72.

[20] Y. Zhang, W. Liu, W. Lou, and Y. Fang, "Location-based compromisetolerant security mechanisms for wireless sensor networks," *IEEE J. Sel. Areas Commun., Special Issue on Security in Wireless Ad Hoc Networks*, vol. 24, no. 2, pp. 247–260, Feb. 2006.

[21] L. Ma, X. Cheng, F. Liu, F. An, and M. Rivera, "iPAK: An in-situ pairwise key bootstrapping scheme for wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 8, pp. 1174–1184, Aug. 2007.

[22] R. Zhang, Y. Zhang, and K. Ren, "DP2AC: Distributed privacypreserving access control in sensor networks," in *IEEE INFOCOM'09*, Rio de Janeiro, Brazil, Apr. 2009.

[23] R. Zhang, J. Shi, Y. Liu, and Y. Zhang, "Verifiable fine-grained top-k queries in tiered sensor networks," in *INFOCOM'10*, San Diego, CA, Mar. 2010.

[24] P. Rogaway, M. Bellare, and J. Black, "OCB: A block-cipher mode of operation for efficient authenticated encryption," *ACM Trans. Inf. Syst. Secure.*, vol. 6, no. 3, pp. 365–403, Aug. 2003.

[25] D. Liu and P. Ning, "Multilevel μTESLA: Broadcast authentication for distributed sensor networks," *Trans. Embedded Computing Sys.*, vol. 3, no. 4, pp. 800–836, 2004.

IJERT