Volume 4, Issue 6, June 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Privacy and Integrity Preservation of Equality Tuple Queries using pairing based homomorphic cryptosystem in WSN

Snehalata k Funde*, Prof. A.D. Gujar
Computer Engineering
&Pune University,
India

Abstract— In today's world the need of WSN has increased not only in critical areas like military applications but also it has drastically increased in day today life like weather information and agriculture field. As WSN's gives an efficient and robust platform for all applications which will otherwise be hard to implement its use is widespread and diverse. Due to these reasons privacy and integrity of WSN data is of great concern. In this paper we do extensive study of the existing methods for preserving privacy and integrity. We also propose and implement a new system which provides privacy preservation and integrity to equality tuple queries. The system has an added benefit that it reduces the overall time taken in transactions and hence reducing the overall cost.

Keywords— Integrity; Privacy; Queries; Tuple; WSN

I. INTRODUCTION

In early days sensor nodes were being used for collecting data from remote places but collecting data from different sensors was very time consuming task. There was memory limitation for every sensor node. So in next few years storage node is added as an interface in between user and sensor node for transferring data in efficient manner. Due to inclusion of storage node in wireless sensor network data from different remote sensors can be sent and collected to nearest storage node and user can access that data anytime easily by passing query very easily when needed it. As wireless sensor networks incur low cost and are efficient enough for military as well as civilian applications, their usage is increasing rapidly in all streams of life. Some of the applications are surveillance, tracking at critical centers, monitoring animal habitats and weather monitoring centers. Attacking the data transferred through these nodes is a cakewalk for any attacker. Most of the times sensitive information travels in between these nodes especially in case of military applications so there security aspect comes in view. Out of the three major bottlenecks namely storage, power and security the former two have been already dealt to a large extent with efficiency, but security still remains a concern. Physical attacks are also a major concern in this field. As WSN's are being used widely there is a need for more emphasis on their security, even though many techniques have been in use and are proposed till date but the degree of sensitivity of data increases the need for more secure measures. In some cases cryptography is not enough so some researchers propose a more secure model sensor trust model, but the summary of referring to security issue is that there is a need to put more emphasis on security. Wireless communication had picked up additional popularity in later years. The requisition driven constrain behind the popularity is simple arrangement and versatility. Moreover the wide requisitions of remote neighborhood organize today; rising provisions of wireless communication incorporate remote sensor systems and Mesh Networks [6]. It might be effectively seen that remote systems administration will pick up additional popularity and tremendous data will be carried on remote systems in the close future. Then again, remote correspondence media is a show media, which represents an enormous test of how to secure data running on the system. Regardless of solid encryption of the information, remote correspondence media still uncovered some data about the activity carried on the system. This is a natural symptom of remote correspondence. Portability implies that the correspondence is normal all over in the sending zone, which in this manner uncovered the correspondence to conceivable ambushers. Simple sending implies that there is sure openness in the convention, which 1 accordingly uncovered some convention data to conceivable ambushers. Area protection is an essential security issue. Misfortune of area protection can empower resulting introduction of personality data on the grounds that area data empowers tying between the internet data and physical world elements. Case in point, web surfing bundles leaving a home in a Cross section system empower a busybody to break down the surfing propensities of one family if the source area of the aforementioned bundles might be resolved. In a remote sensor system, area data frequently implies the physical area of the occasion, which is critical provided for them a few requisitions of remote sensor systems. For case, in a theater of operations, the area of an officer ought not to be uncovered when he starts a show question. In the panda hunter issue, the area of the panda ought not to be uncovered to seekers. A remote sensor system might be a low obligation cycle system. Frequently, movement has an in number association with a certain occasion around then. This gives enormous focal points to a meddler since it doesn't need complex systems to segregate activity around diverse occasions.

II. RELATED WORK

Sheng B, et al. investigates the issue in the setting of a system increased with storage nodes and focus at range queries. It uses bucketing plan to mix the information for range, utilization message encryption for information trustworthiness, and utilize encoding numbers to prevent storage node from dropping information or data[1]. Shij, et al. presents a novel spatio-temporal crosscheck methodology to guarantee secure range queries in event driven two-level sensor systems. It offers information privacy by avoiding master node from taking information and additionally empowers proficient range query handling. All the more essentially, it permits the system possessor to check with quite high likelihood if a query outcome is valid and finish by inspecting the spatial and transient relationships around the returned information[4]. Fei C et al. proposes SafeQ, a protocol that anticipates third party from retrieving data from both sensors gathered information and sink issued queries. SafeQ likewise permits a sink to recognize bargained storage node when they act mischievously. To protect security, SafeQ utilizes a novel strategy to encode both data and queries such that a storage node can effectively process encoded queries over encoded information without knowing their qualities [2]. To safeguard respectability, it proposes another data structure called neighbourhood chains that permit a sink to confirm if the after effect of a query holds precisely the information things that fulfill the query. Furthermore, it proposes a solution for adjusting SafeQ for event driven sensor systems [5]. Rui Z et al. presents three plans whereby the system holder can confirm the credibility and culmination of fine-grained top- k query brings about tired sensor networks, which is the first work of its benevolent. The proposed plans are based symmetric cryptographic primitives and power bargained master node to return both authentic and complete top-k query results to abstain from being gotten [3]. G. Tsudik et al. addresses protection dangers in database outsourcing situations where confide in the administration supplier is restricted. It proposes a model for the basic security utility trade off and outlines a novel calculation for attaining the coveted harmony between protection and utility of the list.[7] R. Agrawal et al. plan has been intended to be sent in requisition situations in which the gatecrasher can get access to the encoded database, yet does not have earlier area data, for example, the appropriation of qualities and can't encode or unscramble discretionary qualities of his decision. The encryption is vigorous against estimation of the accurate esteem in such situations[8]. D. X. Song et al. depicts our cryptographic plans for the issue of seeking on encoded information and give verifications of security for the ensuing crypto frameworks[9]. P. Golle et al. characterizes a security model for conjunctive catchphrase seek over encoded information and present the first plans for directing such looks safely and proposes first a plan for which the correspondence expense is direct in the amount of reports, however that cost might be acquired "disconnected from the net" before the conjunctive question is asked. The security of this plan depends on the Decisional Diffie-Hellman (DDH) presumption. D. Boneh et al. develops open key frameworks that help correlation questions ($x \geq an$) on encoded data and in addition more general queries, for example, subset queries ($x \in S$).These frameworks help self-assertive conjunctive queries (P1^… pi) without releasing data on unique conjuncts. H. String et al. had made a plan for clients with check that their query effects are finished and true. The plan backings range determination on key and non-key characteristics, extend and also join queries on social databases. M.narasimhaand et al. stretches out the plan to furnish both credibility and culmination certifications of queries answers and examines the new approach for different base query sorts and contrasts it and verified Data Structures [10]. W. Cheng et al. proposed a methodology to include confirmation data into a spatial data structure, by developing ensured chains on the focuses inside each one segment, and in addition on all the parts in the data space[11]. Given a query, it creates confirmation that each data point inside those interims of the certified chains that cover the query window either is returned accordingly esteem, or neglects to meet some query condition.

III . EXISTING SYSTEM AND TECHNIQUES

3.1 Paillers Cryptosystem

The plan is an added substance homomorphic cryptosystem; this implies that, given just the public-key and the encryption of m1 and m2, one can process the encryption of m1+m2. It not nly provides privacy to queries in our scheme but also it is the only technique which provides mathematical functions and operators a facility to function freely even in encrypted form thus providing us with a vital edge in our scheme.

3.2 Diffie Hellman Algorithm

Diffie–hellman key exchange (D–h)[nb 1] is a particular system for trading cryptographic keys. It is one of the soonest commonsense cases of key trade executed inside the field of cryptography. The Diffie–hellman key exchange strategy permits two gatherings that have no earlier information of one another to mutually create an imparted mystery key over an unreliable correspondences channel. This key can then be utilized to encode consequent interchanges utilizing a symmetric key. In our scheme it is used when storage node of sensor network and sink exchanges key.

3.3 Prefix numericalization

In Wireless Sensor Network sensor collected data is encrypted individually. So while accessing data from main storage node there is need to process encrypted data as it is without decrypting. Prefix numericalization provides great solution to process data in encrypted form. In Fig. total processing of data is shown. Suppose if we have to check 5 is in between 3 and 7 or not. Then first convert that number into binary form and then apply hash function to it if there is match found in these two sets then 5 is in between 3 and 7.

3.4. Merkle hash

treesIn cryptography and computer science a hash tree or Merkle tree is a tree in which each non-leaf hub is labeled with the hash of the marks of its youngster's nodes. Hash trees are handy since they permit proficient and secure check of the

substance of bigger data structures. Hash trees are a generalization of hash records and hash chains. Exhibiting that a leaf hub is a piece of the given hash tree obliges preparing a measure of data relative to the logarithm of the amount of nodes of the tree;[1] this diverges from hash records, where the sum is corresponding to the amount of nodes. The notion is named after Ralph Merkle. In our scheme it performs integrity check before the user receives the actual data.

3.5Pairing based cryptography

Pairing-based cryptography is the utilization of a pairing between components of two cryptographic assemblies to a third aggregation to develop cryptographic frameworks. Assuming that the same gathering is utilized for the first two aggregations, the pairing is called symmetric and is a mapping from two components of one assembly to a component from a second aggregation. Along these lines, pairings could be utilized to diminish a hard issue in one aggregation to an alternate, normally less demanding issue in an alternate gathering .In our scheme this technique is used for key generation and as well as to speed up the whole operation.

IV . ORIGINAL WORK

4.1 Proposed System Architecture

In our scheme we encrypt equality tuple queries which are sent from sink, the above Fig. 2 gives a brief idea about the whole process. Firstly the user sends equality tuple queries in our approach privacy preservation of these queries is the foremost priority. So, as the user sends these queries they will be received by the pairing based cryptography block. In this block a key is generated and it also helps optimize the speed of the overall system (In a single transaction). Once the key is generated we send the key and the queries to pailler cryptosystem. In this part the overall tuple value i.e. ranges data, time, and node id. Everything will be encrypted. Now, this encrypted equality tuple query will be received at the storage node of sensor networks. At this point of time the storage node uses prefix numericalization method to process the received query, once this process is completed it sends the encrypted data to the user. At the user end again the received encrypted data is firstly verified for integrity using Merkle hash trees and then data is decrypted. The key is exchanged between user and sensor network using Diffie-Hellman algorithm before the encrypted data arrives to the user, the key is for the HMAC function using which the data was initially encrypted.

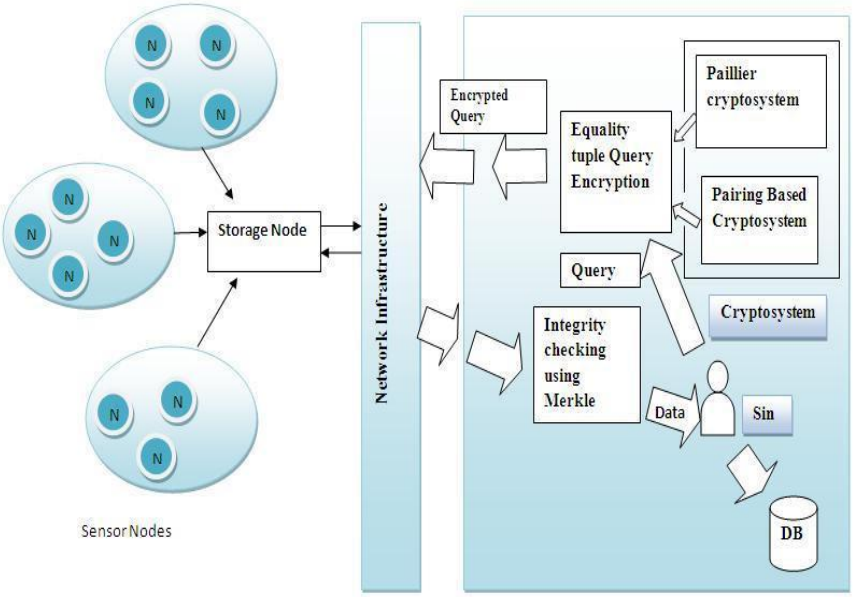

Fig. 1 Proposed System

4.2. Problem Definition

Privacy and integrity preservation of tuple queries using pairing based cryptosystem in WSN.

4.3. Problem Formation

Let N be the total number of nodes in WSN. Let n i be single sensor node where i = {1, 2, 3….N}
Let {n i, t, (d1, d2, dn)} be the data tuple sent by node ni, at time t where (d1, d2….dn) is the data.
We formulate our problem as for sink S, given query {t, (di, dj)}. Sink should be able to preserve privacy of query but time t as well.
Mathematical Formulae Used:
*
1.ø (N) = (p-1) (q-1) =| ZN |
λ (N)=lcm(p-1,q-1)
|Z*N2|=ø (N2) =N ø (N)
∀ x ∈ Z*N2 the following relations are true
X ⋌ (N) ≡ 1 MOD N
X N⋌ (N) ≡ 1 MOD N

4.4. Algorithm:

Step 1: Node ni uses key generation method for encrypting time t field of tuple {t, (di, dj)}.
Key generation: Let k be the security parameter, choose two random k –bit prime numbers p and q and set N =pq.
Choose random base g Є B.

Step 2: Node n encrypt time t using following method,
Node n i chooses a random value r Є ZN and computes the cipher text as
C= g m r N mod N2 and send tuple {C, (di, dj)}.

Step 3: User sends query tuple {t, (di, dj)} where t is the field which is searched on set encrypted value C i where i = {1, 2,3…n} and data values are searched using prefix numericalization method.

Step4: Encrypted data E (D) received at the user end is decrypted at the user end by using public key P k which is exchanged using Diffie-Hellman key exchange algorithm. The key is used to decrypt the data which has been encrypted using HMAC function.

Step5: Decrypted Data D (E (D)) is then checked for integrity using merkle hash tree and data is delivered to the user

V . RESULT

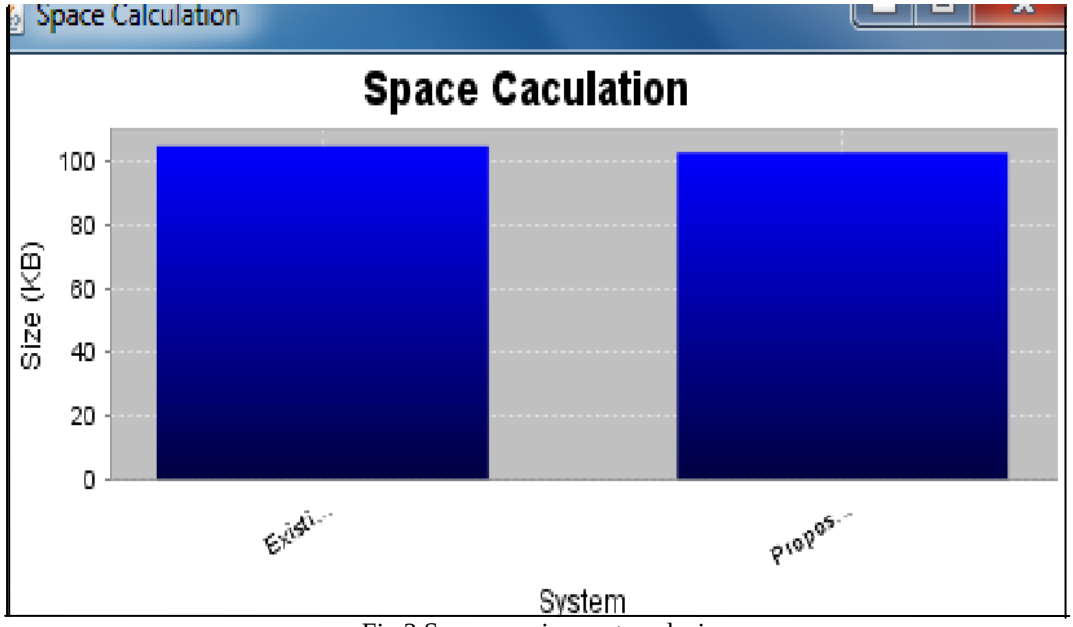

Fig.2 Space requirement analysis

Sample Data set for our work taken from intel lab is given below:-

2004-03-01 00:01:57.13085 5648 1 18.4498 43.1191 43.24 2.67532
2004-03-01 00:02:50.458234 5650 1 18.44 43.0858 43.24 2.66332
2004-03-01 00:04:26.606602 5653 1 18.44 43.1191 43.24 2.65143
2004-03-01 00:05:28.379208 5655 1 18.4498 43.0524 43.24 2.65143
2004-03-01 00:05:50.456126 5656 1 18.4302 43.1525 43.24 2.66332
2004-03-01 00:09:26.66726 5663 1 18.44 43.1858 43.24 2.66332
2004-03-01 00:09:50.555614 5664 1 18.4302 43.2525 43.24 2.65143
2004-03-01 00:01:57.13085 5648 1 18.4498 43.1191 43.24 2.67532

Sample query from sink to storage node is given below:-

2004-03-01,00:01:57-00:04:26,5648-5653,18.4498-18.46,43.1191-43.1524,43.24-43.24,2.66332-2.67532
2004-03-01,00:01:57-00:05:50,5648-5656,18.44-18.4498,43.0524-43.1525,43.24-43.24,2.65143-67532
2004-03-01,00:01:57-00:04:26,>5648,<18.4498,>43.1191,>43.24,>2.66332

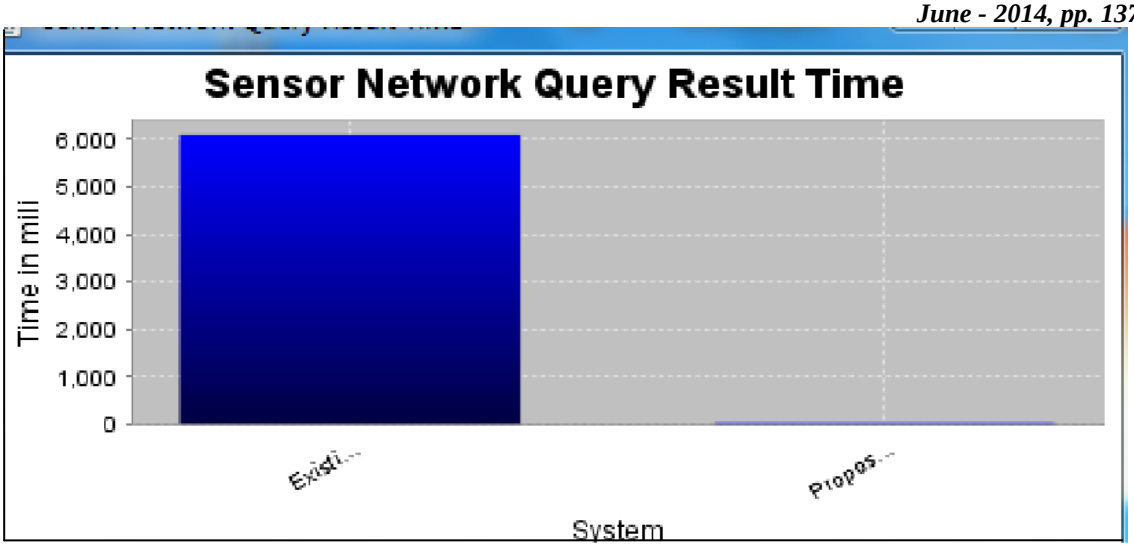

Fig.3 Time requirement analysis

In this section we have provided the sample data set we have considered from intel lab it is three dimensional data. Also sample query fired from sink is also provided. The results in figure 2 suggest that the space requirement when compared to previous methods is less in our method. Also the time requirement is optimized and reduced drastically. This reduction in space and time is possible due to homomorphic properties of pailler cryptosystem which provide pairing based cryptography.

VI . CONCLUSION

By far in our method we tried to provide privacy, integrity and reduce the time taken in a single transaction and our initial result shows that we succeeded in that. As in today's world use and need of WSN is increasing also the complexity and necessity is also increasing to provide security. The very same reason motivated us take on the project and we also added the feature of reducing the time taken in single transaction as an additional benefit. In future we can expand the same method for joint queries and triggers as well which will increase the complexity but will help the field of WSN on a large extent due to the increased variety of security provided.

ACKNOWLEDGMENT

I would like to extend my sincere thanks to my guide Prof. A.D. Gujar for his guidance, encouragement and continuous support throughout the course of this work. I also would like to thank all the staff members for their unconditional support. Last but not the least I would like to thank my family members for their everlasting love, support and belief in me throughout my studies and life.

REFERENCES

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

[2] B.Sheng,Q. Li and W.Mao,"Data Storage placement in Sensor Networks",in Proc.ACM MobiHoc,2006,pp.344-355.

[3] B.Sheng,C.C.Tan,Q.Li and W.Mao,"An Approximation Algorithm for Data Storage Placement in Sensor Network",in Proc.WASA,2007,pp. 71-78.

[4] B.Sheng and Q.Li,"Verifiable Privacy-Preserving Range Query in Two tired Sensor Networks",in Proc. IEEE INFOCOM,2008,pp. 46-50.

[5] D.X.Song,D.Wagner, and A. Perring,"Practical Techniques for searches on encrypted data",in Proc. IEEE S&P,2000,pp. 44-55.

[6] Fei Chen, Alex X. Liu "Privacy and Integrity Preserving Range Queries in Sensor Networks" Publication in Networking,IEEE/ACM Transaction on(volume:20,issue:6),Dec 2012,Pages 1774-1787,ISSN:1063-6692.

[7] F.Chen and A.X.Liu,"SafeQ:Secure and Efficient Query Processing in Sensor Networks",in Proc. IEEE INFOCOM, 2010, pp. 1-9

[8] Na Li[a,*] ,Nan Zhang[b],Sajal K.Das[a,] Bhavani Thuraisingham[c],"Privacy Preservation in Wireless Sensor Networks:A state-of-the-art Survey",Ad Hoc Networks 7(2009)1501-1514.

[9] P. Golle,J. Staddon, and B. Waters,"Secure conjunctive keyword search over encrypted data",in Proc. ACNS,2004,pp.31-45.

[10] R. Agrawal,J. Kiernan, R. Srikant, and Y.Xu,"Order preserving Encryption for Numeric Data",in Proc. ACM SIGMOD,2004,pp. 563-574.

[11] W. Cheng, H.Pang and K.L. Tan,"Authenticating Multi-dimensional query results in data publishing",in Proc. DBSec,2006,pp. 60-73.

