SECURE MESSAGE AUTHENTICATED DATA AGGREGATION SCHEME FOR DATABASE-AS-SERVICE

G.Prabhu[1],K.A.Dhamotharan[2]

***Abstract*----Data Mining or knowledge discovery, is the computer-assisted process of digging through and analyzing enormous sets of data and then extracting the meaning of the data. Data mining tools predict behaviors and future trends, allowing businesses to make proactive, knowledge-driven decisions. Data mining derives its name from the similarities between searching for valuable information in a large database. The Existing work presented the Concealed Data Aggregation (CDA) scheme extended from homomorphic public encryption system. It designed for a multi-application environment and the base station extracts application-specific data from aggregated cipher texts. An impact of compromising attacks in single application environments through the construction of multiple groups. The adversary forge data only in compromised groups is not in the whole system. It designed for secure counting capability and degrades damage from unauthorized aggregations.**

The proposed work presented the Secured Message Authenticated Data (SMAD) Aggregation scheme for Database-AS-Service which provides exact-match query functionality for aggregated data and matching for range queries. The searchable encryption scheme used in DAS is used to increase the query-processing efficiency and provide privacy and authenticity of client data. Establish trusted data base server for client data storage. Aggregation of client queries for multiple applications is made with message authentication code encryptions. Client query responsive data are extracted from trusted data server with authenticated concealment. It handles query aggregation security for DAS Model and minimizes the computation cost due to client query aggregates. Uncompromised secret keys improve the client query response for multiple groups.

***Index Terms*---Data aggregation,Authentication**

1. INTRODUCTION

A wireless sensor network (WSN) consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, pressure, etc. and to cooperatively pass their data through the network to a main location. Wireless sensor networks (WSNs) enable new applications and require non-conventional paradigms for protocol design due to several constraints. Owing to the requirement for low device complexity together with low energy consumption (i.e., long network lifetime), a proper balance between communication and signal/data processing capabilities must be found. This motivates a huge effort in research activities, standardization process, and industrial investments on this field since the last decade.

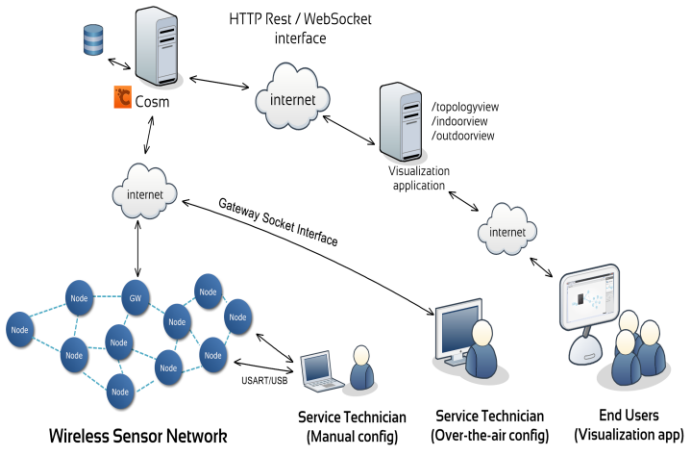

Fig 1. Wireless Sensor Networks

ISSN: 2278 – 7798

A wireless ad hoc network is a decentralized type of wireless network. The network is ad hoc because it does not rely on a pre existing infrastructure, such as routers in wired networks or access points in managed (infrastructure) wireless networks. An ad hoc network typically refers to any set of networks where all devices have equal status on a network and are free to associate with any other ad hoc network device in link range.

Data Mining, also known as Knowledge-Discovery in Databases (KDD), is the process of automatically searching large volumes of data for patterns. Data Mining applies many older computational techniques from statistics, machine learning and pattern recognition.

Data mining consists of five major elements :

- ✓ Extract, transform, and load transaction data onto the data warehouse system.
- ✓ Store and manage the data in a multidimensional database system.
- ✓ Provide data access to business analysts and information technology professionals.
- ✓ Analyze the data by application software.
- ✓ Present the data in a useful format, such as a graph or table.

Data aggregation scheme reduces the large amount of transmission in Wireless Sensor Networks (WSN). Homomorphic encryptions applied to conceal communication during aggregation.

Enciphered data aggregated algebraically without decryption that adversaries are not able to forge aggregated results by compromise. Homomorphic encryption suffers from not satisfy multi-application environments. Insecure on compromised sensor nodes are suffer from unauthorized aggregation attacks.

2. LITERATURE SURVEY

Security is one of the major issues for the ubiquitous sensor networks' [1] (USNs) applications. The evolution of ubiquitous sensor networks provides a unique solution for many ubiquitous information services. Apparently, it challenges the data security and secrecy due to its hostile deployment which is not robust to physical attacks from restricted sources. In order to resolve, the security issue that are duly required in sensor networks, the cryptography protocol is implemented at sensor nodes for node-to-node encryption,considering the data redundancy, energy constraint and security requirement. In this paper, we analyze secure data aggregation scheme and compare its performance with upgraded results of secure protocol called Dragon-Mac, which supports node-to-node encryption using Dragon algorithm based on secrecy methods in sensor networks. This procedure utilizes the entity verification and message authentication through the performance

ISSN: 2278 – 7798

of authenticated encryption scheme in Telos B wireless sensor nodes.

In a large sensor network, in network dat a aggregation [2] significantly reduces the amount of communication and energy consumption. Recently, the research community has proposed a robust aggregation framework called synopsis diffusion which combines multipath routing schemes with duplicate-insensitive algorithms to accurately compute aggregates (e.g., predicate Count, Sum) in spite of message losses resulting from node and transmission failures. However, this aggregation framework does not address the problem of false subaggregate values contributed by compromised nodes resulting in large errors in the aggregate computed at the base station, which is the root node in the aggregation hierarchy. This is an important problem since sensor networks are highly vulnerable to node compromises due to the unattended nature of sensor nodes and the lack of tamper-resistant hardware.

In unattended and hostile environments [3], node compromise can become a disastrous threat to wireless sensor networks and introduce uncertainty in the aggregation results.A compromised node often tends to completely reveal its secrets to the adversary which in turn renders purely cryptography-based approaches vulnerable. How to secure the information aggregation pro cess against compromised-node attacks and quantify the uncertainty existing in the aggregation results has become an important research issue. In this paper, we address this problem by proposing a trust based framework, which is rooted in sound statistics and some other distinct and yet closely coupled techniques. The trustworthiness (reputation) of each individual sensor node is evaluated by using an information theoretic concept, Kullback-Leibler (KL) distance, to identify the compromised nodes through an unsupervised learning algorithm. Upon aggregating, an opinion, a metric of the degree of belief, is generated to represent the uncertainty in the aggregati on result.

In-network data aggregation plays an important role [4] to reduce the transmitting energy and data redundancy in large scale sensor network s. However, the security of wireless sensor networks is a challenging problem in the process of data aggregatio n.An efficient secure data aggregation is proposed to enhance the data security of wireless sensor networks. Firstly, the secure in-network aggregation tree is introduced, then a judgment method based on trust schema was used to detect whether a sensor node has potential misbehavior. After the detection, a local recovery schema is proposed to reduce the possibility to become isolated nodes, which will increase the security level in data aggregation in wireless sensor netw orks. The performance analysis shows the efficiency of the schema.

In-network data aggregation [5] is an essential technique in mission critical wireless sensor networks (WSNs) for achieving effective transmission and hence better power conservation. Common security protocols for aggregated WSNs are either hop-by-hop

ISSN: 2278 – 7798

or end-to-end, each of which has its own encryption schemes considering different security primitives. End-to-end encrypted data aggregation protocols introduce maximum data secrecy with in-efficient data aggregation and more vulnerability to active attacks, while hop by hop data aggregation proto cols introduce maximum data integrity with efficient data aggregation and more vulnerability to passive attacks. A secure aggregation protocol for aggregated WSNs deployed in hostile environments in which dual attack modes are presents. It is a blend of flexible data aggregation as in hop-by-hop protocols and optimal data confidentiality as in end -to-end protocols. Our protocol introduces an efficient O(1) heuristic for checking data integrity along with cost-effective heuristic-based divide and conquer attestation process which is O(ln n) in average -O(n) in the worst scenario- for further verification of aggregated results

Energy is a scarce resource in Wireless Sensor Networks [6]. Some studies show that more than 70% of energy is consumed in data transmission. Since most of the time, the sensed information is redundant due to geographically collocated sensors, most of this energy can be saved through data aggregation. Furthermore, data aggregation improves bandwidth usage. Unfortunately, while aggregation eliminates redundancy it makes data integrity verification more complicated since the received data is unique. A new protocol provides secure aggregation for wireless sensor networks . Our protocol is based on a two hops verification mechanism of data integrity. Our solution is essentially different from existing solutions in that it does not require referring to the base station for verifying and detecting faulty aggregated readings, thus providing a totally distributed scheme to guarantee data integrity. We carried out simulations using Tiny OS environment. Simulation results show that the proposed protocol yields significant savings in energy consumption while preserving data integrity.

Data aggregation is one of the most important [7] techniques in wireless sensor networks to save energy through reducing lots of transmission. However, plaintext aggregation is insecure since eavesdropping or modifying messages is possible.Due to this, concealed data aggregation schemes based on homomor phic encryption have been proposed. Ciphertexts can be operated algebraic computations without decryption in those schemes. Unfortunately, they only provide data confidentiality. While compromising secret in captured sensor nodes, an adversary can still create forged ciphertexts. In this paper, we combines Boneh et al.'s aggregate signature scheme and Mykletun et al.'s concealed data aggregation scheme to overcome the above problems. The proposed scheme aggregates not only ciphertexts but also signatures. Through verifying aggregated signature, data inte grity of each plaintext can be guaranteed..

Data aggregation is a technique [8] used to conserve battery power in wireless sen sor networks (WSN).When securing such a network, it is important that we minimize the number of computationally

ISSN: 2278 – 7798

expensive security operations without compromising on the security. This talk deals with the test-bed implementation of our end to end secure data aggregation algorithm. Unlike previous algorithms which required separate phases for secure aggregation and integrity verification, ours does not require an additional phase for verification. This saves energy by avoiding additional transmissions and computation overhead on the sensor nodes. To sum up our objectives as follows:

1. An encryption algorithm for data confidentiality which will allow us to aggregate encrypted data in a wireless sensor network

2. An aggregate digital signature algorithm to preserve data integrity which allows us to aggregate digital signatures.

3. An implementation of these algorithms which is efficient in computational and communicational aspects in energy constrained environments.

A wireless sensor network (WSN) [9] is a collection of a large number of sensor nodes that have limited computation, communication and power resources. Due to the limited resources, the amount of data transmission should be minimized such that the lifetime of the sensor nodes and bandwidth utilization of the network can be improved. Due to this, the concept of data aggregation has come into the picture. Data aggregation is the process of combining the data coming from various sources and enroute them after removing redundancy such as to improve the overall network lifetime. The in-network processing is done on the aggregator node. The aggregator node aggregate the data received from its child node as per the required aggregation function (like min, max, average, sum etc.) and send the aggregated result to the other high level aggregated node or sink. But in hostile environment these aggregated result should be protected from the various type of attacks in order to achieve data confidentiality, data integrity and source authentication. So security is necessary to be employed with data aggregation.

Wireless sensor networks are vulnerable [10] to many types of security attacks, including false data injection, data forgery and eavesdropping. Sensor nodes can be compromised by intruders and the compromised nodes can distort data integrity by injecting false data. The transmission of false data depletes the constrained battery power and degrades the bandwidth utilization. False data can be injected by compromised sensor nodes in various ways, including data aggregation and relaying. Some sensor nodes are selected dynamically as data aggregators and the nodes between two consecutive data aggregators are called forwarding nodes simply because they forward data. To detect false data injected by a data aggregator while performing data aggregation, some neighboring nodes of the data aggregator (called monitoring nodes) also perform data aggregation and compute MACs for the aggregated data to enable their pair mates to verify the data later.

This project presents a Data Aggregation and Authentication protocol, called DAA, to integrate false data detection

ISSN: 2278 – 7798

with data aggregation and confidentiality.

3. SECURED MESSAGE AUTHENTICATED DATA AGGREGATION SCHEME FOR DATABASE-AS-SERVICE

The phases involved in the proposed scheme are:

- Trusted Database Service Model
- Searchable Message Authentication Code

3.1 TRUSTED DATABASE SERVICE MODEL

In DAS with trusted server which receives queries with tags for the data computed itself. It provides security guarantees for the server not to learn anything about data of user which beyond its occurrence profile unable to identify user access pattern. Adversary cannot mount a chosen-plaintext attack after seeing database or otherwise obtain a priori information about data other than message space.

It is highly undesirable to correlate all the places in the plaintext which occurs semantic correlation with other attribute values. As for authentication of cipher texts gives guarantees integrity in any modification or substitution to the encrypted data is detected by the user. Authentication is ensured at the field level not in record level for entire database. Protect against non-adversarial transmission or storage errors in the database.

3.2 SEARCHABLE MESSAGE AUTHENTICATION CODE

In searchable message authentication code has encrypt attribute values of database. It allows efficient processing of exact-match queries on encrypted aggregated data. Searchable MAC is a standard symmetric encryption which gives message authentication code. It define new symmetric encryption scheme for MACs uses hash based MAC (HMAC). Hash function is collision-resistant resulting in Mac-and-encrypt is secure. It uses a random one-to-one mapping whose output is included with a cipher text to facilitate searchability that the map need not be random, or even pseudorandom,

Security of HMAC relies on non-standard hash also has ENCRYPT-WITH-MAC which present construction has more computation-efficient on client side and more communication-efficient over the network. When the users have low bandwidth connection to database or are connecting via a battery-constrained device they may use Mac of plaintext inside encryption as randomness used in encryption algorithm.

ISSN: 2278 – 7798

4. EXPERIMENTAL RESULTS AND DISCUSSIONS OF SECURED MESSAGE AUTHENTICATED DATA AGGREGATION SCHEME FOR DATABASE-AS-SERVICES

In this section we evaluate performance of Secured Message Authenticated Data (SMAD) Aggregation Scheme for Database-As-Services. One of the major contributions of this work is to provides an exact-match query functionality for aggregated data. To confirm the analytical results, we implemented Secured Message Authenticated Data (SMAD) Aggregation scheme in Java and evaluated the performance of techniques.

The performance of Secured Message Authenticated Data (SMAD) Aggregation scheme is evaluated by the following metrics.

- Communication Overhead
- Computation Overhead
- Mixed data aggregation Cost

Table 1. Communication Overhead

Group size	Existing CDA	Proposed SMAD
10	7	4
20	16	8
30	24	13
40	32	16

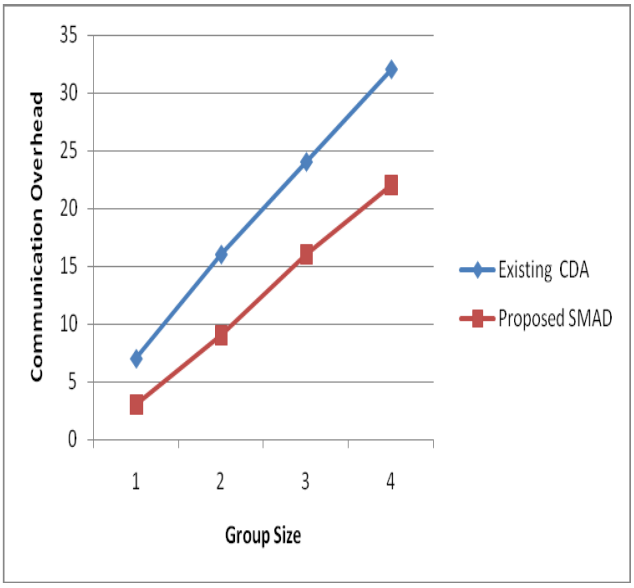

Fig 1. Communication Overhead

Figure1 demonstrates the communication overhead. X axis represents group size whereas Y axis denotes communication overhead using both the existing CDA and proposed SMAD scheme. When the number of group size increased, Communication Overhead gets decreases accordingly. The rate of communication overhead is illustrated using the existing CDA and proposed SMAD Scheme. Figure 1 shows better performance of Secured Message Authenticated Data (SMAD) Aggregation scheme in terms of group size than the existing CDA. SMAD achieves 20 to 35% less Communication Overhead variation when compared to existing system.

ISSN: 2278 – 7798

Table 2. Computation Overhead

Group size	Existing CDA	Proposed SMAD
10	15	9
20	22	12
30	29	19
40	34	26

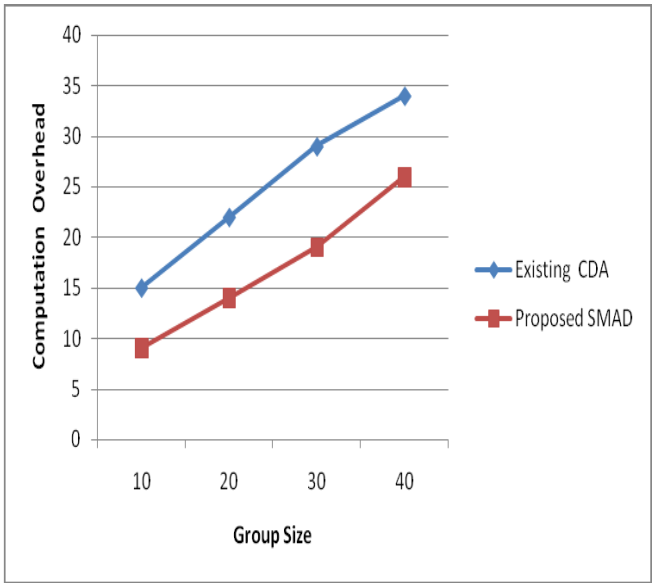

Fig 2. Computation Overhead

Figure 2 demonstrates the computation overhead. X axis represents group size whereas Y axis denotes computation overhead using both the existing CDA and proposed SMAD scheme. When the number of group size increased, Computation Overhead gets decreases accordingly. The rate of computation overhead is illustrated using the existing CDA and proposed SMAD Scheme. Figure 2 shows better performance of Secured Message Authenticated Data (SMAD) Aggregation scheme in terms of group size than the existing CDA.SMAD achieves 20 to 35% less computation Overhead variation when compared to existing system.

Table 3. Mixed Data Aggregation Cost

Group size	Existing CDA	Proposed SMAD
10	8	5
20	17	12
30	26	19
40	35	28

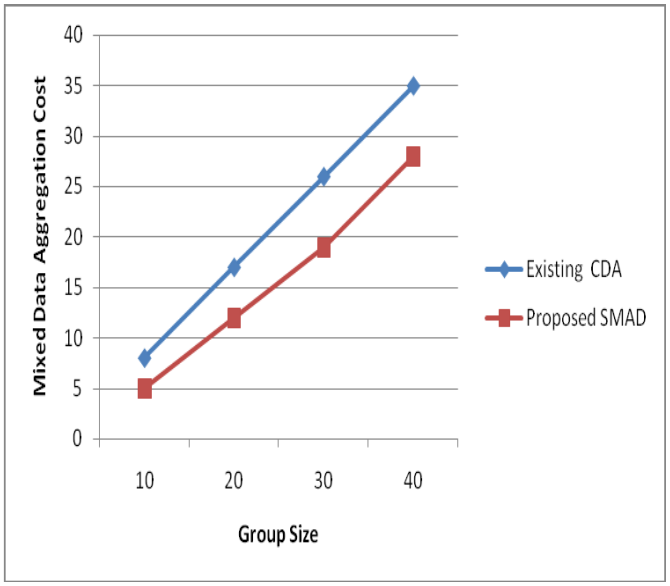

Fig 3. Mixed Data Aggregation cost

ISSN: 2278 – 7798

Figure 3 demonstrates the Mixed Data Aggregation Cost. X axis represents group size whereas Y axis denotes Mixed Data Aggregation Cost using both the existing CDA and proposed SMAD scheme. When the number of group size increased, Mixed Data Aggregation Cost gets decreases accordingly. The rate of Mixed Data Aggregation Cost is illustrated using the existing CDA and proposed SMAD Scheme. Figure 3 shows better performance of Secured Message Authenticated Data (SMAD) Aggregation scheme in terms of group size than the existing CDA.SMAD achieves 20 to 35% less Mixed Data Aggregation Cost variation when compared to existing system.

5. CONCLUSION

The problem of secure data aggregation focuses on enhancing the data availability and the accuracy of the aggregated data. By monitoring neighborhood's activities, each sensor node evaluates the behavior of its cell members in order to filter out the inconsistent data in the presence of multiple compromised nodes. Data aggregation protocols avoid the transmission of redundant data from the sensor nodes. To make the data transmission and aggregation more secured cluster-head is not required to decrypt or encrypt the data received from the sensor nodes. The symmetric keys that are used due to their low memory space and computing requirements, are not transmitted between the cluster-head and the sensor nodes.

REFERENCES

[1] Hoon Jae Lee "A Secure Data Mechanism for Ubiquitous Sensor Network with Dragon Cipher" Published in:INC, IMS and IDC, 2009.

[2] Conti, M. ; Setia, S. ; Jajodia, S. "Secure Data Aggregation in Wireless Sensor Networks" Published in: Information Forensics and Security.

[3] Das, S.K. ; Yonghe Liu "A Trust Based Framework for Secure Data Aggregation in Wireless Sensor Networks" Published in: Sensor and Ad Hoc Communications and Networks, 2006. SECON '06. 2006

[4] Guanghui Li ; Guoying Wang "Efficient Secure In-Network Data Aggregation in Wireless Sensor Networks" Published in:Networks Security Wireless Communications and Trusted Computing (NSWCTC), 2010

ISSN: 2278 – 7798

[5] Aly, S.A. "Secure hop-by-hop aggregation of end-to-end concealed data in wireless sensor networks" INFOCOM Workshops 2008,

[6] Lasla, N. ; Ouadjaout, A. ; Challal, Y. "SEDAN: Secure and Efficient protocol for Data Aggregation in wireless sensor Networks" Local Computer Networks, 2007. LCN 2007.

[7]Yue-Hsun Lin ; Ying-Chu Hsiao ; Chien-Ming Chen "An Efficient and Verifiable Concealed Data Aggregation Scheme in Wireless Sensor Networks " Embedded Software and Systems, 2008.

[8]Madria, S.K. "Secure data aggregation and collaboration in wireless sensor networks" Collaboration Technologies and Systems (CTS), 2011

[9] Mukesh Kumar Jha and T. P. Sharma "A New Approach to Secure Data Aggregation protocol for Wireless Sensor Network" (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 05, 2010, 1539-1543

[10] S. Chaitanya Rami and Reddy, P. Ravinder Kumar "Implementation of Data Aggregation and Authentication in Wireless Sensor Networks" International Journal of Soft computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-5, November 2012

AUTHORS

G.Prabhu PG Scholar, Erode Sengunthar Engineering College, Thudupathi, Erode.

K.A.Dhamotharan Assistant Professor Sr.G, Erode Sengunthar Engineering College, Thudupathi, Erode.

ISSN: 2278 – 7798