Data Integrity Preserved Data Aggregation Technique over Smart grid Communication System

Saranya Alagesan, Mrs.P.Anatha Prabha

***Abstract*— Smart grid application is the most concerned application in the real world environment which will generate different power readings in different time periods. This need to be gathered and send to the centralized server for further processing. The single node failure in the smart grid system might lead to entire system failure where the aggregation cannot be performed well. This problem is resolved in the existing work by using the fault tolerance based data aggregation technique where the data can be aggregated even if any of the nodes failed in the system. However data integration becomes the greatest issue in the smart grid system where the data forwarded to the centralized server might get changed in case of corruption. This problem is resolved in the existing work by introducing the recoverable scheme in which data can be recovered even in case of corruption also. This is achieved by comparing with the average aggregated value with all the data that are sensed by the sensors.**

***Index Terms*— dataintegrity, sensors, privacy preserving, aggregation, smart grid**

I. INTRODUCTION

Wireless Sensor Network refers to the group of spatially dispersed and dedicated sensors for monitoring and recording the physical conditions of the environment and organising the collected data at a central location. To enabling the power distribution to be more effective and consistent from power generation, transmission, and distribution to customers consumption, and supports the renewable energy . By deploying various sensors along the two-way flows, i.e., the electricity flow and the communication flow, a huge amount of real-time information is reported and collected to the control center (CC) for timely monitoring and analyzing the health of the power grid. Smart meters (SMs) are important components of smart grid. They are two-way communication devices deployed at consumers premise, records power consumption sporadically. With smart meters, smart grid is able to collect real time information about grid operations and eminence at an operation center, through a reliable communications network deployed in parallel to the power transmission and distribution grid. It is responsible for dynamically adjusting power supply to meet demand, and sensing and replying to weaknesses or failures in the power system in real time. However, the real-time usage data, e.g., collected every 30 seconds, contain personalized power usage patterns, which are highly relevant to users' privacy; thus, they must be endangered from illegal entities. Up to now, many data aggregation schemes have been proposed to preserve individual user privacy in smart grid.Most of them use the homomorphic encryption techniques to encrypt user's data, so that the *semi-trust* aggregator (e.g., the GW) can aggregate all users'data without decryption. However, they only consider the protection of users' privacy against the GW(aggregator), while the CC, if considered under the *honest-but-curious* model, is still able to learn individual user's data, as the keys owned by the CC may be not only utilized to decrypt the gathered data, but also used to reveal any user's electricity usage.

However, the solutions proposed in adopt the "honest-but-curious" model to assume that all the smart meters follow the protocol properly. Although it protects data privacy against curious smart meters, it does not consider accidental errors or cyber-attacks that tamper with the protocol. Therefore, it is vulnerable to unintentional errors (e.g. accidental errors in network transmission, storage and computing) and compromised meters or communication channels. For instance, a malfunctioning meter may accidentally produce errors in computing the aggregation; a compromised meter may drop intermediate aggregation results, and submit a random value to its parent node; an adversary who has hijacked the connection between two meters may external adversaries tamper with the aggregation process, expecting to mess up with load balancing, resource allocation and smart pricing. To protect data integrity against accidental errors, we first introduce an end-to-end authentication scheme that is compatible with the homomorphic encryption based in-network aggregation schemes proposed.

In particular, a homomorphic signature is generated for the aggregated metering data at each intermediate node along with the aggregation process. In the end, the collector could effectively verify the correctness of the aggregation by checking the consistency between the aggregation result and the aggregation signature. The homomorphic signature scheme needs no decryption and re-encryption at in-between meters, to facilitate an efficient signing/verification process. Moreover, to defend against fake data injection attacks, we present a hop-by-hop signature and incremental verification scheme. In this solution, aggregated outputs from smart meters are signed, and signatures are managed in a distributed manner (instead of transmitted to the collector on-the-fly). Verification is only performed ,when anomalies in the aggregation results are detected at the collector. The incremental verification process efficiently traces the anomaly in a breath-first manner, which is computationally inexpensive. More importantly, it ensures faithfulness and undeniability properties, so that the faulty nodes are always identified with undeniable evidences.

Saranya Alagesan, sri krishna college of technology, coimbatore, india.

Assistant Prof.Mrs.P.Anatha Prabha, Department Of Computer Science and Engineering, sri krishna college of technology, coimbatore, india

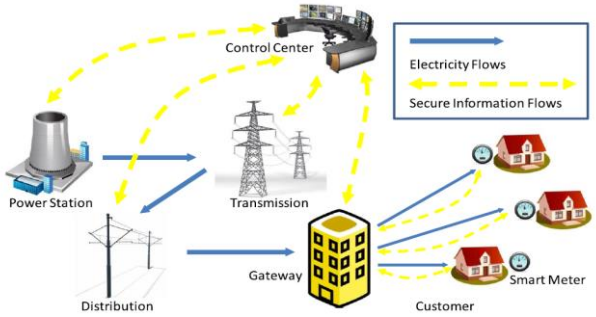

Fig1.smart grid system architecture

II. Related Work

Instead of using costly trusted third party to anonymize the metering data, a more efficient approach is to hide individual data via aggregation[5][6]. To proposed a no leakage protocol to aggregate partial shares of smart meter readings in a neighborhood using an additively homomorphic encryption scheme[5] . However, the approach is not scalable due to the high communication overhead. with the privacy-preserving in network data aggregation[1][3]. Differing from the wireless sensor network approaches that focus on defending against misinformation, the in-network aggregation solutions in smart metering aim to protect end-to-end data confidentiality and privacy against malicious or "curious" meters en route, but neglect authentication mechanisms for data integrity protection[4][7]. To address the problem, simple authentication schemes based on consensual PKI digital signature scheme or cryptographic MAC have been proposed[8]. However, they are either not compatible or introduce excessive hop-by-hop verification overhead[2][4]. Therefore, we present a new homomorphic signature based authentication scheme that can efficiently re-generate signatures for aggregation results at intermediate meters but also support batch verification at the collector[2][6]. Homomorphic signature scheme was first proposed in to authenticate packets in network coding protocols and later extended to applications as delegatable data sharing and data outsourcing .

III. Problem Formalization

In this section we formalize system model,security requirements and the design goal.

3.1System model

In our system model, we primarily focus on how to report residential users privacy-preserving electricity usage data to the operation center in smart grid communications[6][7]. Specifically,we consider a typical residential area (RA), which comprises a local gateway connected with smart grid operation center, and a large number of residential usersU1; U2; . . . ; Uwg[4][9],. The system has two types of events involved: a set of *users* and an *aggregator*. we consider a typical smart grid communication design for residential users, which includes a trusted authority (TA), a CC(control center), a residential gateway (GW), and a great number of residential n users in residential area(RA).

3.2Security Requirements

Security is important for the success of protected smart grid communications[5][9].However, there exists an adversary A residing in the RA to eavesdrop the residential users reports. More seriously, the adversary A could also intrude in the database of the GW and the smart grid operation center to steal the individual user reports.

Confidentiality

Where an adversary may compromise the privacy of residential users by eavesdropping the communication data from the residential users to the GW and those from the GW to the CC[3]. In such a way, each individual user's electricity usage data can achieve the privacy-preserving requirement[5][4]. In addition, the confidentiality requirement also includes the OA's(operational authority) responses should be privacy-preserving, i.e., only the legal residential users in the RA can read them[3][6].

Data integrity

Authenticating an encrypted report that is really sent by a legal residential user and has not been altered during the transmission, i.e., if the adversary A forgoes and/or modifies a report, the malicious operations should be detected[2]. Where an adversary is usually the participants of the protocol including the GW or the CC,which could access or misuse the information of residential users to compromise their privacy, or the curious residential users, who actively seek or infer other users'private usage data[7][9].

3.3 Design goal

Under the aforesaid system model and security requirements, our design goal is to develop an efficient and integrity on privacy-preserving aggregation scheme for secure smart grid communications[10].

*The computation of data integrity should be permitted in the proposed aggregation scheme***:**

In order to provide diversified service for users, the CC may need to determine integration of users data[7][9].To preserve user privacy, those functions should be computed in the form of ciphertext[7][8]. Therefore, the proposed system should allow the GW to compute to integrate the data aggregation without decryption.

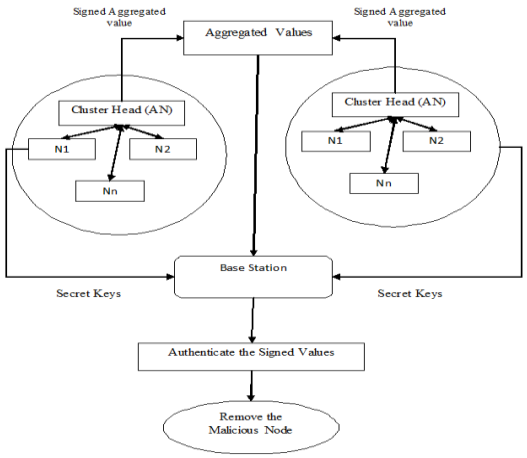

Fig 2 architecture diagram

IV. PRELIMINARIES

4.1.Homomorphic Encryption

Homomorphic encryption is employed in to support aggregation operations on concealed data, so that data privacy is well protected from intermediate meters[8][4]. The solution in focuses on data confidentiality and privacy, but lacks the capability to verify data integrity[6][3]. Conventionally, authentication and integrity check are supported by appending digital signatures to the data. However, due to the malleability property of homomorphism, homomorphic encryption based schemes do not provide non-repudiation and thus cannot support verification of individual inputs at either intermediate meters or the final destination (i.e., the collector). Therefore, additional techniques are needed for signing multiparty metering data and evaluating the integrity of the aggregation results.

1)Key generation: The key generation algorithm remains largely the same to select secret key sk = a 2 Zq for all smart meters and private key pk = ga as its public key for verifiers.get the tuple (p, q,G), where p, q aredifferent primes with$|p| = |q| = \tau$, and G is a cyclicgroup of order $N = pq$. Randomly chose two generators $g, x \in G$ and set $h = xq$. Then, h is a random generator of the subgroup of G having order p. Eventually, the public and private keys are $PK = (N,G, g, h)$ and $SK = p$,respectively.

2) Encryption: Given a message $m \in \{0, 1, \ldots , V \}$, where $V << q$ is the bound of the message space, choose a random number $r \in ZN$ then, the ciphertext can be calculated as $C = gmhr \in G$.

3) Decryption: Given the private key $SK = p$ and the Cipher text $C \in G$, first compute $Cp = (gmhr)p = (gp)m$. Let $gp = gp$, then $Cp = gmp$. To recover m, it comes down to compute the discrete logarithm of gmp .

Note that when m is a short message, say $m \leq V$ for some small bound V, the decryption takes expected time$O(\sqrt{V})$ utilizing Pollard's lambda method. The Boneh–Goh–Nissim cryptosystem has additive homomorphism property.

Suppose $C1, C2 \in G$ be two cipher texts for messages $m1,m2 \in \{0, 1, \ldots , V \}$, respectively, to obtain the cipher texts of $m1 + m2$, one can simply compute the product $C = C1C2hr$ for a random $r \in ZN$.

4.2Signing:

For smart meter Ni (whose unique identifier is IDi), let Coi 2 Zq be the encrypted form of the plaintextoutput Poi 2 Zq after homomorphic encryption for in-network aggregation. Ni computes hi = H(IDi) and outputs thesignature _aggi = (hiCoi)a 2 G for < IDi;Coi >.

Ni and Idi	smart meter in the NAN and its unique identifier
Pi and Poi	input and aggregation output of Ni in plaintext
Ci and Coi	homomorphic encrypted form of Pi and Poi

V. PROPOSED METHOD

In the proposed work, we introduce a concept named Recoverable Data integrity. In RCDI, a base station can recover each sensing data produced by all sensors even if these data have been aggregated by cluster heads (aggregators). With these distinct data, two functionalities are provided. First, the base station can verify the integrity and authenticity of all detecting data. Second, the base station can execute any aggregation functions on them. Then, we propose two RCDI schemes named RCDI-HOMO and RCDI-HETE for homogeneous and heterogeneous WSN respectively. In the security analysis, we determine that the future schemes are secure under our attack model.

RCDI-HOMO is composed of four techniques: Setup, Encrypt-Sign, Aggregate, and Verify. The Setup procedure is to prepare and install necessary secrets for the CC and each sensor. When a sensor decides to send sensing data to its CH, it performs Encrypt-Sign and sends the effect to the CH. Once the CH receives all results from its members, it activates Aggregate to aggregate what it established, and then refers the final results (aggregated ciphertext and signature) to the BS. The last procedure is Verify. The CC first extracts individual sensing data by decrypting the aggregated ciphertext. Afterward, the CC verifies the authenticity and integrity of the decrypted data based on the corresponding aggregated signature.

5.1 Proposed RCDI-schemes

A RCDI SCHEME FOR HOMOGENEOUS WSN (RCDI-HOMO) In this section, we propose a recoverable concealed data integration scheme named RCDI-HOMO for homogeneous WSN.Lists the notations that we will use later. Construction of RCDI-HOMO RCDI-HOMO is composed of four measures: Setup, Encrypt-Sign, Aggregate, and Verify. The Setup procedure is to arrange and install needed secrets for the BS and each sensor. When a sensor decides to send sensing data to its CH, it performs Encrypt-Sign and sends the effect to the CH. Once the CH accepts all results from its members, it activates Aggregate to aggregate what it established, and then sends the final results (aggregated ciphertext and signature) to the BS. The last procedure is Verify. The BS first extracts distinct sensing data by decrypting the aggregated ciphertext.

RCDI-HETE Scheme Here, we challenge to fully abuse H-Sensors which have stronger computing capability. Operations on L-Sensors could be switched to H-Sensors. In addition, H-Sensors can be calculated to be tamper-resistant, so we may allow HSensors to store the restricted secret information if required. With these concerns, we redesign an RCDI scheme named RCDI-HETE. While the use of tamper-resistant devices may upgrade the hardware cost; however, in a heterogeneous WSN, majority of sensors are low-end sensors (L-Sensors). In our design, totalling cost on L-Sensors is switched to H-Sensors, so L-Sensors can be very low-priced and unpretentious. In fact, the overall hardware cost is reduced. RCDI-HETE is composed of five measures: Setup, Intracluster Encrypt, Intercluster Encrypt, integrate, and Verify. In the Setup procedure, necessary secrets are burdened to each H-Sensor and L-Sensor. Intracluster Encrypt procedure involves when L-Sensors desire to send their sensing facts to the corresponding H-Sensor. In the Intercluster Encrypt procedure, each H-Sensor aggregates the received data and then encrypts and signs the integrate result.

In addition, if an H-Sensor receives ciphertexts and signatures from other H-Sensors on its defeating path, it activates the Aggregate procedure. Finally, the Verify procedure ensures the reality and integrity of each combined result

5.2 Recovery Property

The Recovery property attempts to provide two techniques. First, BS can verify the integrity and authenticity of all sensing data. Second, BS can perform random aggregation operations on these data. However, in RCDI-HETE, the BS only recovers distinct aggregated result generated by each cluster rather than all sensing data. Now we will show that RCDI-HETE also provides these functionalities. RCDI-HETE can verify each sensing data through the aid of H-Sensors. More precisely, Intracluster Encrypt method allows L-Sensor Lj i to send not only Ekj, but also the MAC (message authentication code) of to its cluster head Hj; therefore, Hj can verify the integrity of the data sent from its cluster members. 2. Every H-Sensor is loaded several necessary aggregation functions before deployment, so the BS can command every H-Sensor to perform the designated aggregation function. For example, if BS decides to obtain the summation of all data, it assigns HSensors to perform the addition operation. Then, the BS can perform thelast addition when it recovers every result from every H-Sensor. Also, if BS then decides to perform maximum-selection operation, the BS notifies every H-Sensor to select the determined value between the sensing data in the Intercluster Encrypt procedure.

VI. Security and scalability analysis

In this section, we demonstrate the proposed schemes are secure under the attack model defined .More detailed security analysis and scalability analysis are described of the Supplemental Material available online. We first assume that an adversary does not concession sensors. The future schemes are secure because sensing messages are encrypted. In RCDI-HOMO, each sensor encrypts their messages with PBS before transmitting. In RCDI-HETE, intracluster traffic is encrypted with pairwise keys. Further, our scheme generates the corresponding signature for each sensing data. Consequently, an adversary cannot modify messages and introduce fake messages since he cannot sign forged messages without private keys. If an rival has the ability to concession sensors, we consider the following situations.

An adversary can compromise a sensor and perform it as a authorized one. Sensing compromised sensors that still act normally is infeasible in all existing detection mechanisms in WSN. Also, if the value adversary can also try to manipulate the aggregated result. He may generate false data, modify authorized messages, or imitate other sensors. The proposed schemes are still secure against above attacks because of the signature needed for each produced message. On the other hand, we discuss the situation when an adversary compromises a cluster head in RCDI-HOMO. First, he cannot decrypt the aggregated ciphertext or each individual ciphertext because no decryption private key is kept in a cluster. Second, the compromised cluster head may selectively drop some ciphertexts and signatures in the Combined procedure. This kind of attack which is called selective forwarding attack. Fortunately, previous research proposed mechanisms to defend against this attack.

VII. Performance and Cost Evaluation

To calculate the performance of the proposed schemes, performance time (or “delay”) is the main measurement of performance evaluation. Without loss of overview, we define dispensation delay and aggregation delay for deployed sensors. Processing interruption indicates the execution time for sensors to produce ciphertexts and corresponding signatures before transmission. Aggregation delay is also calculated by measuring time spent on processing time on aggregating ciphertexts and signatures in the proposed schemes. The last suspension, decryption delay, is not considered since the base station is considerably prevailing as a workstation. Therefore, this delay is insignificant and can be ignored. Another criterion is cost evaluation. Cost estimation involves communication and calculation aspects.

The last scheme, RCDI-HETE, has been revised from naı¨ve RCDI-HETE to enhance the performance of LSensors. Processing delay on L-Sensors decreases (2.97 ms) since Intraencrypt controls symmetric cryptography. Most of calculation cost has been switch to H-Sensors instead. Although Interencrypt is related to Encrypt-Sign, H-Sensors executes better and saves more energy than L-Sensors. Another improvement is the decreased communication costs. To summarize the results from the proposed schemes, RCDI-HETE utilizes the benefits and advantages of HSensors. The naı¨ve RCDI-HETE reduces the corresponding delays during aggregations compared with RCDI-HOMO. However, H-Sensors require more energy on communication in naı¨ve RCDI-HETE. we further simulate a WSN while applying RCDI-HOMO, RCDI-HETE, and Nonaggregate model.

According to the above comparisons, RCDI-HOMO seems to be the worst in performance evaluation. This is because RCDI-HOMO provides better security. Fortunately, the overall cost in RCDI-HOMO is still affordable for WSN. On the other hand, CDA and scheme could association other secure mechanisms to achieve the same security level with RCDI-HOMO. However, the cost of elaborate mechanisms is high and volatile.

VIII. Conclusion

In this paper, we have proposed recoverable concealed data integration schemes for homogeneous/heterogeneous WSNs. A special feature is that the base station can securely recover all sensing data rather than integrated results, but the transmission overhead is still acceptable. Moreover, we integrate the aggregate signature scheme to ensure data authenticity and integrity in the design. Even though signatures bring additional costs, the proposed schemes are still affordable for WSNs after evaluation. Considering a large WSN (over 100 nodes), we also performed simulations on the proposed schemes.

REFERENCES

[1] S. Ozdemir and Y. Xiao, "Secure data aggregation in wireless sensor networks: A comprehensive overview," Comput. Netw., vol. 53, no. 12, pp. 2022–2037, Aug. 2009.

[2] L. Wasserman, All of statistics : a concise course in statistical inference. New York: Springer.

[3] C. T. Chou, A. Ignatovic, and W. Hu, "Efficient computation of robust average of compressive sensing data in wireless sensor networks in the presence of sensor faults," Parallel and Distributed Systems, IEEE Transactions on, vol. 24, no. 8, pp. 1525–1534, Aug 2013.

[4] R. Lu, X. Liang, X. Li, X. Lin, and X. Shen,"EPPA: An efficient andprivacy-preserving aggregation scheme for secure smart grid communications,"*IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1621–1631,Sep. 2012.

[5] M. Hashmi, S. Hanninen, and K. Maki, "Survey of smart grid concepts,architectures, and technological demonstrations worldwide," in IEEE PES Conference on Innovative Smart Grid Technologies, 2011.

[6] X. Li *et al.*, "Securing smart grid: Cyber attacks, countermeasures, and challenges," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 38–45, Aug. 2012

[7] L. Chen, R. Lu, and Z. Cao, "PDAFT: A privacy-preserving data aggregation scheme with fault tolerance for smart grid communications,"*Peer-to-Peer Netw. Appl.*, to be published.

[8] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart gridsusing homomorphic encryption," in *Proc. 1st IEEE Int. Conf. Smart GridCommun. (SmartGridComm)*, 2010, pp. 327–332.

[9] A. Metke and R. Ekl, "Smart grid security technology," in *Innovative Smart Grid Technologies*, Jan. 2010, pp. 1–7.

[10] Efthymiou C, Kalogridis G (2010) Smart grid privacy via anonymization of smart metering data. In: SmartGridComm, pp 238–243

[11]Jia W, Zhu H, Cao Z, Dong X, Xiao C Human-factor-awareprivacy preserving aggregation in smart grid. IEEE Syst J (toappear)

[12]T.-H. Chan, E. Shi, and D. Song, "Privacy-preserving stream aggregationwith fault tolerance," in *Financial Cryptography and Data Security*,A. Keromytis, Ed. Berlin Heidelberg: Springer, 2012,LectureNotes in Computer Science, vol. 7397, pp. 200–214.

[13].HULL, B., BYCHKOVSKY, V., ZHANG, Y., CHEN, K., GORACZKO,M., MIU, A., SHIH, E., BALAKRISHNAN, H., AND MADDEN, S.Cartel: a distributed mobile sensor computing system. In *ACM*
SenSys (2006).

[14]. A. Wander, N. Gura, H. Eberle, V. Gupta, and S. Shantz, "Energy Analysis of Public-Key Cryptography for Wireless Sensor Networks," Proc. IEEE Third Int'l Conf. Pervasive Computing and Comm. (PERCOM '06), 2005.

[15] G. De Meulenaer, F. Gosset, F.X. Standaert, and L. Vandendorpe, "On the Energy Cost of Communication and Cryptography in Wireless Sensor Networks," Proc. IEEE Int'l Conf. Wireless and Mobile Computing, Networking and Comm., pp. 580-585, 2008.