Hindawi
Wireless Communications and Mobile Computing
Volume 2017, Article ID 2539673, 8 pages
https://doi.org/10.1155/2017/2539673

Research Article

Privacy-Preserving Meter Report Protocol of Isolated Smart Grid Devices

Zhiwei Wang and Hao Xie

School of Computer Sciences, Nanjing University of Posts and Telecommunications, Nanjing, China

Correspondence should be addressed to Zhiwei Wang; zhwwang@njupt.edu.cn

Received 17 January 2017; Revised 24 April 2017; Accepted 4 May 2017; Published 6 June 2017

Academic Editor: Zhe Yang

Smart grid aims to improve the reliability, efficiency, and security of the traditional grid, which allows two-way transmission and efficiency-driven response. However, a main concern of this new technique is that the fine-grained metering data may leak the personal privacy information of the customers. Thus, the data aggregation mechanism for privacy protection is required for the meter report protocol in smart grid. In this paper, we propose an efficient privacy-preserving meter report protocol for the isolated smart grid devices. Our protocol consists of an encryption scheme with additively homomorphic property and a linearly homomorphic signature scheme, where the linearly homomorphic signature scheme is suitable for privacy-preserving data aggregation. We also provide security analysis of our protocol in the context of some typical attacks in smart grid. The implementation of our protocol on the Intel Edison platform shows that our protocol is efficient enough for the physical constrained devices, like smart meters.

1. Introduction

While the swift advances in smart grid are triggering radical innovations in this field, today's power grid is widely different from the traditional grid [1–4]. Traditional grid has the characteristic of centralized one-way transmission, which only transmits electricity from the generation plants to customers. Smart grid is featured with intelligent transmission (decentralized two-way transmission) and distribution networks, which combines the traditional grid and the new information processing technologies. On the one hand, smart grid integrates more green energies such as solar and wind power into energy supply; on the other hand, it improves the reliability, security, and efficiency of electric system by two-way communication of consumption data and other electric system's operations. In general, smart grid can realize the intelligent electricity generation, resource allocation, and dynamic pricing.

In this system, smart grid devices such as smart meters play an important role for collecting the power usage data and the status data. Such data are generated by some plug-in monitor sensors. In general, the smart grid data communication network can be divided into four layers [5] as Figure 1 shows. Various sensors and other smart grid devices consisting of a home area network are the first layer. Then, the smart meters and a neighborhood gateway which form a neighborhood area network are the second layer. Furthermore, all the neighborhood gateways connecting each other consist of the third layer network. Moreover, the forth layer network is a high speed public network through fiber gateways which is responsible for transfer all the data to the data center in electricity service provider (ESP).

However, not all smart grid devices are connected to the smart grid data communication network, due to the network outrage or opt-out agreement between the customers and the ESP. According to the utility-scale smart meter deployments report [6] published by Electric Innovation at Edison Foundation, the smart meters only cover 43% US homes. Some smart grid devices are located sparsely and far away from the data center of ESP. Thus, it would be a heavy cost to extend the smart grid data communication network for covering such isolated smart grid devices. Moreover, some in-network smart grid devices also will be disconnected from the smart grid network due to the natural disasters such as tornado and

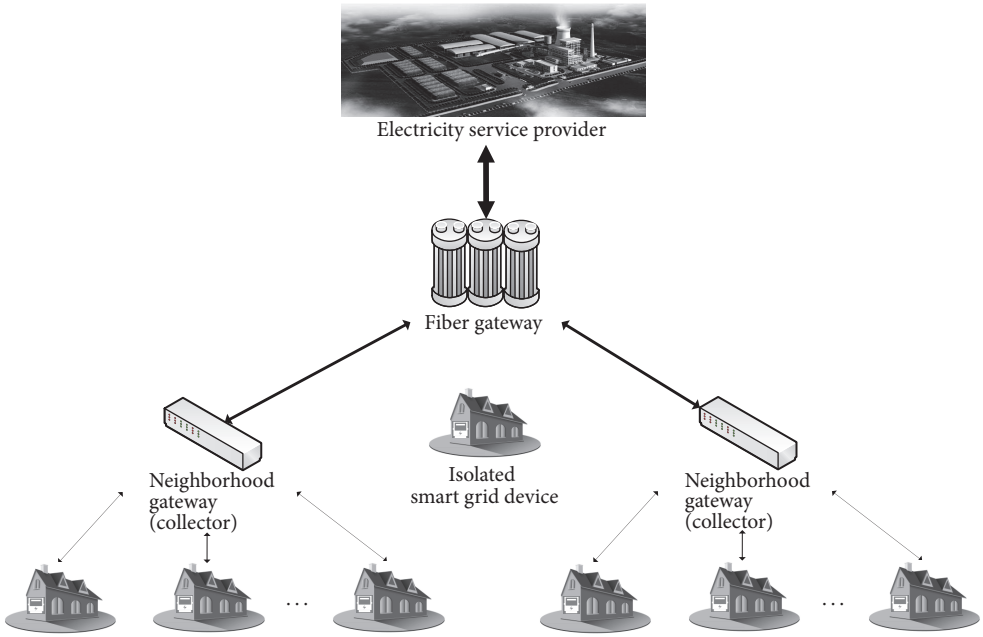

Figure 1: Smart grid data communication network.

earthquake. Thus, for such isolated smart grid devices, the ESP may send a worker to the location of them and read the power usage data by using the handhold smart meter reader.

In general, several protocols are used in smart grid communication network [7], for the propose of authentication, power allocation, meter reporting, and so on. The meter report protocol is used to calculate the total monthly power consumption data for each individual customers. For the isolated smart grid device, a smart reader device should be used as a bridge between the ESP and it as Figure 2 shows. Although the smart reader device needs to read the smart meter more frequently for monitoring the energy supply, the ESP only needs to obtain the total long-term consumption data for the energy forecast.

Up to now, several privacy-protection aggregation schemes have been proposed. Li et al. [8] constructed an incremental aggregation scheme based on a virtual aggregation tree which relies on the topology of network. Garcia and Jacobs [7] proposed an aggregation scheme combined with additive secret sharing. Lu et al. [9] proposed an efficient privacy-preserving scheme for multidimensional data structure. The three schemes are all based on Pallier's homomorphic encryption technology. Fan et al. [10] proposed data aggregations scheme based on the subgroup indistinguishability assumption. All the above aggregation schemes are designed for the in-network smart grid devices, and they are used to aggregate individual usage date from different customers. For the isolate smart grid devices, Sha et al. [5] proposed a secure and efficient authentication protocol, but their meter report protocol did not provide a data aggregation mechanism for privacy-preserving. For the isolated smart grid devices, there exists the same drawback as in-network devices that fine-grained power usage data may leak the personal privacy information [11, 12]. If a corrupted worker in the ESP can obtain the fine-grained power usage data, then he can analyze the daily activities of the customer. Thus, a secure data aggregation mechanism for privacy protection is also required for isolated smart grid devices. The fine-grained power usage data should be protected in the reader device and cannot be leaked to anyone else.

This paper aims to propose an efficient privacy-preserving meter report protocol for the isolate smart grid devices. The protocol not only contains an additively homomorphic encryption scheme used to aggregate the encrypted data but also includes a linearly homomorphic signature scheme [13, 14] for protection against unintentional errors and altering messages in malicious. Furthermore, both the isolated smart grid devices and the reader devices have only restricted resources, and thus both the encryption and signature schemes should provide the high performance in terms of efficiency.

The contributions of this paper can be listed as follows: (1) We propose an encryption scheme with additively homomorphic property to aggregate the encrypted metering data. To be compatible with the data aggregation, we also propose a linearly homomorphic signature scheme which is used to sign the ciphertext of metering data. The signatures will be aggregated along with the ciphertexts stored in the reader device. This allows the ESP to verify the correctness of aggregated result by checking the aggregation signature. (2) We provide a security analysis to our meter report protocol in context of several typical attacks in smart grid. (3) To evaluate the appropriacy of our meter report protocol for the resource-constrained devices, we implement our protocol on the Intel Edison platform which is a development system for Internet of Things (IoT) devices.

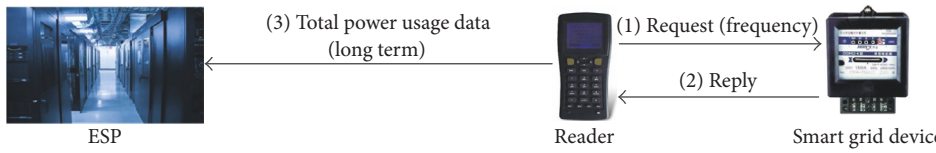

Figure 2: Meter report of isolated smart grid devices.

Organization. Related mathematical concepts to our construction and proofs are reviewed in Section 2. The privacy-preserving meter report protocol for isolated smart grid devices is proposed in Section 3. We analyze our protocol against several typical attacks in Section 4. Section 5 discusses the performance of our protocol on the platform of MacBook Pro and Edison. Finally, we conclude our paper in Section 6.

2. Preliminary

In this section, we review related mathematical concepts for our construction and proofs.

Assuming that G and G_T are two cyclic groups with the prime order p, we define $e : G \times G \to G_T$ to be the bilinear map as it has the following properties:

(1) Bilinear: $\forall g_1, g_2 \in G$, $a_1, a_2 \in Z_p$, $e(g_1^{a_1}, g_2^{a_2}) = e(g_1, g_2)^{a_1 a_2}$.

(2) Nondegenerate: $\exists g \in G, e(g, g) \neq 1$.

(3) Efficient computability: there exists an efficient algorithm to compute $e(g_1, g_2)$ for all $g_1, g_2 \in G$.

We define the q-strong Diffie-Hellman (q-SDH) assumption over G as follows.

Definition 1 (q-SDH assumption). Let $\text{Gen}(1^\iota)$ be a group generation algorithm that takes a security parameter ι as input and outputs a description of a prime order group $\Theta = \{p, G, G_T, e\}$. The q-SDH assumption over group G states that, for any probabilistic polynomial-time (PPT) attackers, given a tuple $(g, g^\beta, g^{\beta^2}, \ldots, g^{\beta^q})$ for randomly chosen $\beta \xrightarrow{R} Z_p$ and $g \xrightarrow{R} G$, the advantage for obtaining a solution $(\gamma, g^{1/(\beta+\gamma)})$ is negligible in ι, where $\gamma \in Z_p$.

Next, we define two composite order groups (G', G'_T) with order $N = pq$, where p and q are distinct large primes. Thus, G is a product of two groups $G' = G'_p \times G'_q$, and their orders are p and q, respectively. In essence, the subgroup indistinguishability assumption is that an element in group G' is computationally indistinguishable from a random element in G'_p or G'_q. Let g' be a generator of G'. We define a nongenerate and efficiently computable bilinear map $e : G' \times G' \to G'_T$ over G' and G'_T. The subgroup indistinguishability assumption [15] can be described as follows.

Definition 2 (subgroup indistinguishability assumption). Let $\text{Gen}(1^\iota)$ be a group generation algorithm that takes a security parameter ι as input and outputs a description of a multiplicative group $\Psi = \{p, q, G', G'_T, e'\}$, where $G' = G'_p \times G'_q$. The subgroup indistinguishability assumption over group G' states that, for any PPT attackers, the advantage

$$\begin{aligned}&\text{Adv}_{\mathbf{A}}(\iota)\\ &= \left|\Pr\left[\mathbf{A}(\Psi, x) = 1; x \longleftarrow_R G'\right] - \Pr\left[\mathbf{A}(\Psi, x^q) = 1\right]\right|,\\ &\text{Adv}_{\mathbf{A}}(\iota)\\ &= \left|\Pr\left[\mathbf{A}(\Psi, x) = 1; x \longleftarrow_R G'\right] - \Pr\left[\mathbf{A}(\Psi, x^p) = 1\right]\right|\end{aligned} \tag{1}$$

is negligible in ι.

3. Design of Meter Report Protocol

3.1. System Model. There are three parties including electricity service provider (ESP), reader, and isolated smart grid device in the system model of the proposed protocol. The ESP and the isolated smart grid device should setup their public/secret key pairs and other public information. When the reader tries to frequently collect the encrypted metering data from the isolated smart grid device, several attacks may be possible. Firstly, an attacker may listen to the communications between the reader and the isolated smart grid device to obtain the metering data or alter the messages. Secondly, a corrupted reader may be used to obtain the power usage data. Thirdly, a corrupted reader may provide an incorrect total power usage data to the ESP. Finally, a fake ESP worker may analyze the power usage data with fine granularity to identify the daily activities of the customer.

In the meter report model as Figure 2 shows, the reader needs to much more frequently read from the smart grid device for monitoring the energy supply. Each time the reader reads, the smart grid device encrypts its metering data with a random number and signs it before he sends it to the reader. After a long term, the ESP can only obtain the total power usage data of the customer.

3.2. Construction. The proposed protocol consists of four phases, which will be described in detail as follows. Some notations can be defined here.

(i) $H : \{0, 1\}^* \to Z_N^*$ is a one-way hash function.

(ii) t is the tag of currently regular period.

(iii) ID_{esp} is the identity information of electricity service provider.

(iv) r_i is the ith random number chosen by smart grid device.

(v) r_0 is the sum of random numbers $\sum r_i$.

(vi) α is the secret key of isolated smart grid device.

(vii) Y is the public key of isolated smart grid device.

(1) Setup Phase

(i) ESP: the ESP randomly chooses two distinct large primes (p, q) and computes the RSA parameter $N = pq$ (example initiation: let $\mathbf{P}$, p, and q be distinct large primes such that $\mathbf{P} = 2pq + 1$. Obviously, $Z_{\mathbf{p}}^*$ is a quadratic residue group with order $N = pq$. $Z_{\mathbf{p}}^*$ can be denoted as $Z_{\mathbf{p}}^* = G_p \times G_q$, where G_p and G_q are both prime order cyclic groups. Gonzalez et al. proved that the subgroup decision assumption over $Z_{\mathbf{p}}^*$ holds if the factoring problem over N is hard). It generates u in group G with order N and produces a generator g of the subgroup G_q. Then, it computes $h = u^q$, which is an element in subgroup G_p. Finally, the ESP publishes the public parameters $\{N, u, h, g, \mathrm{ID}_{\mathrm{esp}}\}$ where $\mathrm{ID}_{\mathrm{esp}}$ is its identity information and keeps $\{p, q\}$ as the secret information.

(ii) Isolated smart grid device: the isolated smart grid device randomly chooses $\alpha \in Z_N^*$ as its secret key and publishes the public key $Y = g^{\alpha}$. Then, let $\mathrm{ID}_{\mathrm{esp}}$ denote the identity of ESP who is the customer's energy supplier.

(2) Reading Phase

(i) Isolated smart grid device: when the reader needs to read the metering data m_i for the ith time in a long term, the isolated smart grid device chooses $r_i \in Z_N^*$ randomly and computes a ciphertext $\mathrm{CT}_i = g^{m_i} h^{r_i}$. We assume that reader reads the metering data n times during such a long term. There is a limitation that $r_0 = \sum_{i=1}^{n} r_i$ should not be a large number. Then, the smart grid device computes a signature

$$\begin{aligned} \sigma_1 &= g^{1/(\alpha + H(\mathrm{ID}_{\mathrm{esp}} \| t))}, \\ \sigma_{2i} &= \left(\mathrm{CT}_i\right)^{1/(\alpha + H(\mathrm{ID}_{\mathrm{esp}} \| t))}, \end{aligned} \tag{2}$$

where t is the tag of currently regular period. Finally, it sends $\{\mathrm{CT}_i, (\sigma_1, \sigma_{2i})\}$ to the reader.

(ii) Reader: after receiving $\{\mathrm{CT}_i, (\sigma_1, \sigma_{2i})\}$, the reader verifies identity of its ESP and the currently long term by checking $e(\sigma_1, Y \cdot g^{H(\mathrm{ID}_{\mathrm{esp}} \| t)}) = e(g, g)$. Here, the reader verifies the smart grid device's first signature component to assure that who is its ESP and to avoid that the customer will make payments for an improper ESP. If the signature σ_1 is true, then the reader stores $\{\mathrm{CT}_i, (\sigma_1, \sigma_{2i})\}$.

(3) Aggregation Phase

(i) Isolated smart grid device: at the end of a long term, the isolated smart grid device encrypts r_0 as $\mathrm{CT}_{r_0} = g^{r_0} h^s$ with a random number $s \in Z_N^*$ and sends it to the reader.

(ii) Reader: after receiving $\mathrm{CT}_{r_0} = g^{r_0} h^s$, the reader needs to aggregate the total power usage data of the isolated smart grid device. We assume that the reader has read the smart grid device n times during this long term, and thus n ciphertext/signature pairs $\{\mathrm{CT}_i, (\sigma_1, \sigma_{2i})\}_{i \in [1,n]}$ have been stored in the reader. Then, the reader computes $\mathrm{CT} = \prod_{i=1}^{n} \mathrm{CT}_i$ and $\sigma_2 = \prod_{i=1}^{n} \sigma_{2i}$, and reports $\{\mathrm{CT}, (\sigma_1, \sigma_2), \mathrm{CT}_{r_0}\}$ to the ESP.

(4) Decryption and Verification Phase

(i) ESP: when the ESP receives $\{\mathrm{CT}, (\sigma_1, \sigma_2), \mathrm{CT}_{r_0}\}$, it firstly verifies its identity information and the currently long term by checking $e(\sigma_1, Y \cdot g^{H(\mathrm{ID}_{\mathrm{esp}} \| t)}) = e(g, g)$ and then computes $W = \mathrm{CT}_{r_0}^p = (g^p)^{r_0}$ and $\widetilde{g} = g^p$. Since r_0 is not a large number, the ESP can compute the discrete log of W on the base of $\widetilde{g}$ by using Pollard's lambda method [16] in polynomial time. Then, the ESP computes $V = \mathrm{CT} \cdot h^{-r_0} = g^{\sum_{i=1}^{n} m_i}$. Since the total power usage data $M = \sum_{i=1}^{n} m_i$ is also not a large number, the ESP can compute the discrete log of V on the base of g. Finally, the ESP computes $\varrho = \sigma_2^p$ and verifies σ_2 by checking $e(\varrho, Y \cdot g^{H(\mathrm{ID}_{\mathrm{esp}} \| t)}) = e(\widetilde{g}, g)^M$.

The correctness of the above formulas can be depicted as follows.

Authentication of Its ESP

$$\begin{aligned} &e\left(\sigma_1, Y \cdot g^{H(\mathrm{ID}_{\mathrm{esp}} \| t)}\right) \\ &\quad = e\left(g^{1/(\alpha + H(\mathrm{ID}_{\mathrm{esp}} \| t))}, g^{\alpha} \cdot g^{H(\mathrm{ID}_{\mathrm{esp}} \| t)}\right) = e\left(g, g\right). \end{aligned} \tag{3}$$

Ciphertext Decryption

$$\begin{aligned} W &= \mathrm{CT}_{r_0}^p = g^{r_0 \cdot p} \cdot h^{s \cdot p} = g^{r_0 \cdot p} \cdot u^{s \cdot p \cdot q} = \left(g^p\right)^{r_0} \cdot 1 \\ &= \widetilde{g}^{r_0}, \\ V &= \mathrm{CT} \cdot h^{-r_0} = g^{\sum_{i=1}^{n} m_i} \cdot h^{\sum_{i=1}^{n} r_i + (-r_0)} = g^{\sum_{i=1}^{n} m_i} \cdot h^0 \\ &= g^{\sum_{i=1}^{n} m_i}. \end{aligned} \tag{4}$$

Aggregate Signature Verification

$$\begin{aligned} \varrho = \sigma_2^p &= \left(\left(g^{\sum_{i=1}^{n} m_i} \cdot h^{\sum_{i=1}^{n} r_i}\right)^{1/(\alpha + H(\mathrm{ID}_{\mathrm{esp}} \| t))}\right)^p \\ &= \left(\left(g^p\right)^{\sum_{i=1}^{n} m_i} \cdot u^{\sum_{i=1}^{n} r_i \cdot p \cdot q}\right)^{1/(\alpha + H(\mathrm{ID}_{\mathrm{esp}} \| t))} \\ &= \left(\left(g^p\right)^{\sum_{i=1}^{n} m_i} \cdot 1\right)^{1/(\alpha + H(\mathrm{ID}_{\mathrm{esp}} \| t))} \\ &= \left(\widetilde{g}^{\sum_{i=1}^{n} m_i}\right)^{1/(\alpha + H(\mathrm{ID}_{\mathrm{esp}} \| t))}. \end{aligned} \tag{5}$$

Thus,

$$e\left(\varrho, Y \cdot g^{H(\mathrm{ID}_{\mathrm{esp}}\|t)}\right) = e\left(\left(\widetilde{g}^{\sum_{i=1}^{n} m_i}\right)^{1/(\alpha+H(\mathrm{ID}_{\mathrm{esp}}\|t))}, g^{\alpha+H(\mathrm{ID}_{\mathrm{esp}}\|t)}\right) = e\left(\widetilde{g}, g\right)^{M}. \quad (6)$$

4. Security Analysis

Our privacy-preserving meter report protocol is proposed not only to prevent the unauthorized parties to read or alter the metering data from the isolated smart grid devices, but also to securely aggregate the fine-grained power usage data in a long term. Here, we show the security properties of our scheme in context of six typical attacks in smart grid.

4.1. Against External Attack. The external attackers can eavesdrop on the communication channels to obtain the unauthorized information. In our protocol, all the metering data are encrypted, which provide strong protection to the external attackers. The proof of Theorem A.2 in Appendix shows that our encryption scheme satisfies the CPA secure under the subgroup indistinguishability assumption. The external attackers also cannot alter a metering data of the isolated smart grid device, since they cannot forge a valid signature. Theorems A.4 and A.5 in Appendix show that our linearly homomorphic signature schemes are unforgeable under the q-SDH assumption and Boneh and Boyen signature.

4.2. Against Smart Grid Device Attack. A smart grid device attack is that a fake smart grid device aims to mimic a legitimate device. In our design, we use the signature technology to prevent a fake smart grid device from authenticating with the reader and ESP. Moreover, a fake smart grid device may want to let the customer to pay for an improper ESP, but our design can also avoid this situation, since the first component of linearly homomorphic signature is a signature of the proper ESP's identity, and its unforgeable security is under Boneh and Boyen signature (the security proof of Theorem A.4 can be seen in Appendix).

4.3. Against Internal (Reader) Attack. An attacker may use a lost legitimate reader to obtain the unauthorized information or maliciously alter total the power usage data of a smart grid device, which is called the internal (reader) attack. In reading phase, the legitimate reader only can verify the signature of device's identity. But the power usage data m_i cannot be recovered from the ciphertext $\mathrm{CT}_i = g^{m_i} h^{r_i}$, since the reader cannot get the ESP's secret key (p, q). In aggregation phase, the reader also cannot decrypts CT_{r_0} to get r_0 and obtains the total power usage data. On the other hand, the linearly homomorphic signature and the encryption of r_0 prevent the reader from altering the total power usage data, since it does not know the secret key α of the isolated smart grid device. The unforgeability of our linearly homomorphic signature scheme has been proved by Theorems A.4 and A.5. The properties of linearly homomorphic signature also protect the correctness and integrity of the total power usage data.

4.4. Against Internal (ESP) Attack. We assume that the legitimate workers of ESP make the malicious attacks. After receiving the ciphertext/signature pair $\{\mathrm{CT}, (\sigma_1, \sigma_2), \mathrm{CT}_{r_0}\}$ from the reader, the ESP can compute $V = \mathrm{CT} \cdot h^{-r_0} = g^{\sum_{i=1}^{n} m_i}$ to recover the total power usage data. However, the ESP cannot decrypt the individual metering data m_i from CT and r_0, since it does not know each corresponding random number r_i.

4.5. Against Man-In-The-Middle Attack. A Man-In-The-Middle attacker aims to mimic the right person to fool one side by using the information from another side. In reader-device and ESP-device authentication, a public key based linearly homomorphic signature scheme is used to authenticate the device's identity and the ciphertexts. It provides the strong defense for the Man-In-The-Middle attacks, since the attacker cannot convince the reader and ESP to accept its public key.

4.6. Against Replay Attack. If an attacker obtains the information between the communication of two sides, then he intercepts the communication and replays the information maliciously, which is called replay attack. In our designing, we use the tag of currently term t to prevent the replay attack from different terms. If the attacker wants to modify t in device's signature for the replay attack, then he should get the device's secret key α. However, it is almost impossible to guess the device's secret key. If an attacker wants to make replay attack in the same period, then it should modify r_0 in ciphertext CT_{r_0} that is also impossible.

5. Performance Analysis

Let P denote the pairing computation cost, E denote the exponent cost, and Mu denote the point multiplication. Table 1 shows the computational complexity of our protocol.

Following the theoretical analysis, we test our scheme on two different platforms, where one is a normal personal computer, and the other is a resource-constrained device. We implement our protocol in C with the pairing based cryptography (PBC) library [17] for the underlying arithmetic and pairing operations. We use the Type-A curves as defined in PBC library for the implementation, since the Type-A curves offers the highest efficiency among all the three types of curves.

The first test machine is MacBook Pro with Intel core i5 CPU (2.5 GHz) running Os X 10.9.3, which RAM is 4 GB. The second test machine is Intel Edison development platform, which is designed to rapidly prototype and produce Internet of Things (IoT) products. Since the isolated smart grid device and reader device are usually resource-constrained devices, we test our protocol on this platform. We use Edison platform with a dual-core, dual-threaded Intel Atom CPU at 500 MHz and 1 GB RAM, running Yocto Linux v1.6.

Table 2 shows the time cost of reading phase for smart grid device and reader. We compute the average value on 100 randomized runs. The time cost of isolated smart grid device is about 0.43 seconds, if our protocol is run over the Edison platform. For the reader, it needs 0.42 seconds to verify the signature, while the protocol is run over the Edison platform. In aggregation phase, the time cost of isolated device is about

Table 1: Computational complexity of our protocol.

Computational complexity	ESP	Reader	Isolated device
Reading phase	Null	$1P + 1\text{Mu} + 1E$	$1\text{Mu} + 4E$
Aggregation phase	Null	$2n\text{Mu}$	$1\text{Mu} + 2E$
Decryption and verification phase	$3P + 3\text{Mu} + 6E$	Null	Null

Table 2: Time cost in reading phase.

Platform	MacBook	Edison
Smart grid device	0.02 s	0.43 s
Reader	0.016 s	0.42 s

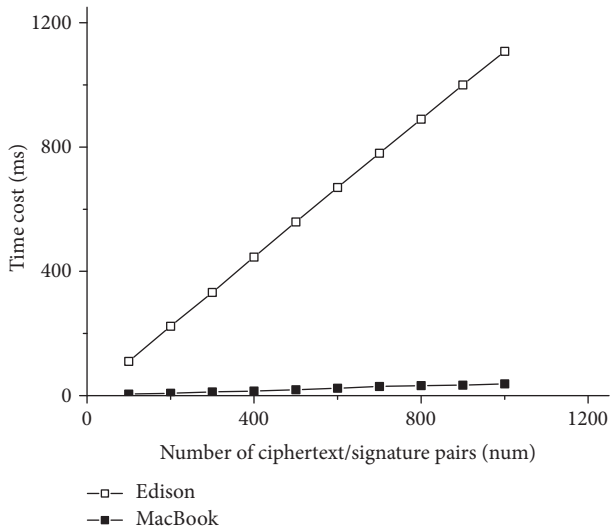

Figure 3: Time Cost of reader in aggregation phase.

1.5 milliseconds on the Edison platform, while it needs about 0.06 milliseconds over MacBook Pro. Figure 3 shows the time cost of reader in aggregation phase. We can see that the time consuming of reader is increased by the number of ciphertext/signature pairs to be aggregated. The time cost of decryption for the ESP is about 77 milliseconds. Although the total power usage data M is increased by the number of individual consumption data m_i, the computation of the discrete log of V is very slightly raised.

6. Conclusion

In practical, the fine-grained individual power consumption data may leak the personal privacy information of the users. Thus, in order to protect the personal privacy, data aggregation mechanism should be designed in the meter report protocol. In this paper, we propose an efficient privacy-preserving meter report protocol for the isolated smart grid devices, which consists of an encryption scheme with additively homomorphic property and a linearly homomorphic signature scheme. To prevent unauthorized seeing the intermediate metering data, the metering data should be encrypted by using the encryption scheme with additively homomorphic property and aggregated using such a property. Besides the encryption scheme, a linearly homomorphic signature scheme which is compatible with data aggregation is also designed in our protocol for verifying the correctness and integrity of the aggregation result. We give security analysis to our protocol in context of six typical attacks in smart grid. The implementation of our protocol on the Edison platform shows that our protocol is efficient enough for the resource-constrained devices.

Appendix

Here, we provide the security proofs to the encryption scheme and the linearly homomorphic signature scheme used in the proposed meter report protocol.

Definition A.1. A public key encryption scheme is CPA secure, if for all the advantage of any PPT attacker A in the following game is negligible in the security parameter ι.

Setup. The challenger obtains the public/secret key pair (pk, sk) by running Setup(1^ι) and sends pk to the attacker, where the public key includes a message space $\mathbf{M}$ and a ciphertext space Ξ. The challenger sets $\text{Enc}(pk, m)$ as an encryption algorithm.

Challenge. The attacker sends two messages m_0 and m_1 with the same length to the challenger. Then, the challenger responds the challenge ciphertext $\text{CT} = \text{Enc}(pk, m_b)$ under a random bit b.

Output. The attacker outputs its guess b' to b. If $b' = b$, then the attacker wins the game.

Theorem A.2. *If the subgroup indistinguishability assumption holds on G, then the above encryption scheme is CPA secure.*

Proof. We assume that there exists an attacker **A** which can break the above encryption scheme with nonnegligible probability $\epsilon(\iota)$ and a challenger **C** that takes an instance of subgroup indistinguishability assumption. We will prove the theorem by an interaction game between **A** and **C**.

Setup. The challenger **C** is given an instance (N, G, G_T, x) of subgroup indistinguishability assumption and generates a generator $g \in G$. Then, it sends the public parameters (N, G, G_T, x, g) to the attacker **A**.

Challenge. **A** chooses two messages m_0 and m_1 with the same length and then sends them to **C**. **C** chooses a random number $r^* \in Z_N^*$ and returns the challenger ciphertext $\text{CT}^* = g^{m_b} x^{r^*}$, where b is a random bit.

Output. **A** outputs its guess b' to b, and if $b' = b$, then **A** wins the game and **C** outputs that "x is uniformly in G_p". Otherwise, **C** outputs that "x is uniformly in G".

If x is uniformly in G, then the challenge ciphertext CT is randomly in G_T, which is independent of b. Thus, $\Pr[b' = b] = 1/2$ in this case. However, if x is uniformly in G_p, then $\Pr[b' = b] = 1/2 + \epsilon(\iota)$ in this case since **A** can break the above encryption scheme with the probability of $\epsilon(\iota)$. The probability difference of these two cases is $\epsilon(\iota)$, which is nonnegligible in our assumption. But it contradicts that the subgroup indistinguishability assumption is hard. Thus, our assumption is not correct, and the encryption scheme is CPA secure. □

Our linearly homomorphic signature scheme is based on Boneh and Boyen signature [18], which has been proved strongly unforgeability against a weak attacker under the q-SDH assumption. Here, we will firstly provide the security definition of linearly homomorphic signature.

Definition A.3. An linearly homomorphic signature scheme is simply unforgeable, if for all the advantage of any PPT attacker **A** in the following game is negligible in the security parameter ι.

Setup. The challenger obtains the public/secret key pair (pk, sk) by running Setup(1^ι) and sends pk to the attacker, where the public key includes a message space **M** and a signature space Σ. The challenger sets Sign as the signing algorithm and Verify as the verification algorithm.

Queries. The attacker sends a random number $x \in \{0,1\}^*$ and a message $m \in \mathbf{M}$ to the challenger for a signature query. Then, if m is the first query for x, the challenger randomly chooses a tag $\tau_x \in Z_N^*$ and gives it to the attacker. Otherwise, the challenger looks up the previously chosen τ_x. The challenger then returns the signature $\sigma \leftarrow \text{Sign}(sk, \tau_x, m)$. This query can be repeated for a polynomial times; however there is a restriction that at most n message can be queried for one tag τ_x. We let V_x denote the set of elements m queried for x.

Output. The attacker outputs a tag $\tau^* \in Z_N^*$, a message $m^* \in \mathbf{M}$, and a signature $\sigma^* \in \Sigma$. The attacker wins if $\text{Verify}(pk, \tau^*, m^*, \sigma^*) = 1$ and satisfies one of the following conditions (the type 2 forgery can be split into 2 subtypes):

Type 1: $\tau^* \neq \tau_x$ for all x queried by attacker (a type 1 forgery).

Type 2: $\tau^* = \tau_x$ for one pair of x, and $m^* \neq \sum_{i=1}^{n} m_i$, where $m_i \in V_x$ (a type 2 forgery).

Type 2(a): the first element σ_1^* of signature (σ_1^*, σ_2^*) output by the attacker is *not* equal to the signature $(\sigma_1, \sigma_2) \leftarrow \text{Sign}(sk, \tau_x, m)$ computed by the challenger.

Type 2(b): the first element σ_1^* of signature (σ_1^*, σ_2^*) output by the attacker equals the signature $(\sigma_1, \sigma_2) \leftarrow \text{Sign}(sk, \tau_x, m)$ computed by the challenger.

The advantage of the attacker is the probability that the attacker wins the game.

We can show that type 1 and type 2(a) forgery in our linearly homomorphic signature scheme will lead to a forgery of the underlying Boneh and Boyen (BB) signature.

Theorem A.4. *Our linearly homomorphic signature scheme is secure against type 1 and type 2(a) forgeries, if BB signature is strong unforgeable against a weak attacker.*

Proof.
Sketch. The challenger simulates the public key of our scheme by using the public key of BB signature and the element $h = g^\delta$. For responding the signature query on m in our scheme, the challenger queries τ to the challenger of BB signature and obtains σ_1. Then, the challenger returns $(\sigma_1, \sigma_2 = \sigma_1^{m+\delta \cdot r})$ for a random number $r \in Z_N^*$. Finally, if the attacker of our scheme outputs a valid forgery (τ^*, m^*, σ^*), then the first component of σ^* is a valid forgery of BB signature. □

Theorem A.5. *Our linearly homomorphic signature scheme is secure against type 2(b) forgeries, if q-SDH assumption holds.*

Proof.
Sketch. The challenger of our scheme takes as input an instance $(g, g^\beta, g^{\beta^2}, \ldots, g^{\beta^q})$ of the q-SDH assumption and forms the polynomial $P(x) = \prod_{i=1}^{q}(x + \tau_i) \in Z_N[t]$ for the q distinct tags $\tau_1, \ldots, \tau_q$ queried by the attacker. Let $P_l(x) = \prod_{i \neq l}(x + \tau_i)$ and $l^* \in \{1, \ldots, q\}$ randomly chosen by the challenger. Then, the challenger constructs $X = g^{P(\alpha)}$, $Y = g^{P_{l^*}(\alpha)}$, and $h = g^\delta$, which can be used to respond to the signature queries from the attacker. Finally, when the attacker returns the forged signature $\sigma^* = (\sigma_1^*, \sigma_2^*)$ on m^* and τ^*, the challenger computes $z = \sigma_2^* / X^{(m^* + \delta \cdot r^*) \cdot 1/(\alpha + \tau^*)}$. If the forged signature is valid, then $z = Y^{1/(\alpha+\tau^*)}$.

Let $P_{l^*}(t)/(t + \tau^*) = Q(t) + c/(t + \tau^*)$, where $Q(t)$ is a polynomial over Z_N^*. Thus, $z = g^{P^*(\alpha)/(\alpha+\tau^*)} = g^{b \cdot (Q(\alpha) + c/(\alpha+\tau^*))}$. Then, $(\tau^*, (z/g^{Q(\alpha)})^{1/c})$ is a solution to the q-SDH assumption. □

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] F. Li, W. Qiao, H. Sun et al., "Smart transmission grid: vision and framework," in *Proceedings of IEEE Transactions on Smart Grid*, vol. 1, pp. 168–177, 2010.

[2] D. Niyato, L. Xiao, and P. Wang, "Machine-to-machine communications for home energy management system in smart grid," *IEEE Communications Magazine*, vol. 49, no. 4, pp. 53–59, 2011.

[3] Z. M. Fadlullah, M. M. Fouda, N. Kato, A. Takeuchi, N. Iwasaki, and Y. Nozaki, "Toward intelligent machine-to-machine communications in smart grid," *IEEE Communications Magazine*, vol. 49, no. 4, pp. 60–65, 2011.

[4] H. Liang, B. J. Choi, W. Zhuang, and X. Shen, "Towards optimal energy store-carry-and-deliver for PHEVs via V2G

system," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM '12)*, pp. 1674–1682, Orlando, Fla, USA, March 2012.

[5] K. Sha, N. Alatrash, and Z. Wang, "A secure and efficient framework to read isolated smart grid devices," *IEEE Transactions on Smart Grid*, 2016.

[6] I. E. I., "Utility-scale smart meter deployments: building block of the evolving power grid," IEI Smart Meter Update, sep 2014.

[7] F. D. Garcia and B. Jacobs, "Privacy-friendly energy-metering via homomorphic encryption," in *Proceedings of 6th International conference Security and Trust Management*, vol. 6710, pp. 226–238, Springer, Berlin, Germany, 2011.

[8] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart grids using homomorphic encryption," in *Proceedings of the 1st IEEE International Conference on Smart Grid Communications (Smart Grid Comm '10)*, pp. 327–332, IEEE, Gaithersburg, Md, USA, October 2010.

[9] R. Lu, X. Liang, X. Li, X. Lin, and X. Shen, "EPPA: an efficient and privacy-preserving aggregation scheme for secure smart grid communications," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1621–1631, 2012.

[10] C.-I. Fan, S.-Y. Huang, and Y.-L. Lai, "Privacy-enhanced data aggregation scheme against internal attackers in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 666–675, 2014.

[11] Z. Fu, K. Ren, J. Shu, X. Sun, and F. Huang, "Enabling personalized search over encrypted outsourced data with efficiency improvement," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 9, pp. 2546–2559, 2015.

[12] T. Ma, J. Zhou, M. Tang et al., "Social network and tag sources based augmenting collaborativerecommender system," *IEICE Transactions on Information and Systems*, vol. E98-D, no. 4, pp. 902–910, 2015.

[13] D. Mandell Freeman, "Improved security for linearly homomorphic signatures: a generic framework," in *Public Key Cryptography – PKC 2012*, pp. 697–714, Springer, Heidelberg, Berlin, Germany, 2012.

[14] Z. Wang, G. Sun, and D. Chen, "A new definition of homomorphic signature for identity management in mobile cloud computing," *Journal of Computer and System Sciences*, vol. 80, no. 3, pp. 546–553, 2014.

[15] Z. Brakerski and S. Goldwasser, "Circular and leakage resilient public-key encryption under subgroup indistinguishability," in *Advances in Cryptology – CRYPTO 2010*, Rabin. and T., Eds., vol. 6223, pp. 1–20, Springer, Heidelberg, Berlin, Germany, 2010.

[16] E. Teske, "Computing discrete logarithms in arithmetic progressions," http://citeseerx.ist.psu.edu/.

[17] B. Lynn, "The pairing-based cryptography (pbc) library," http://crypto.stanford.edu/pbc.

[18] D. Boneh and X. Boyen, "Short signatures without random oracles and the SDH assumption in bilinear groups," *Journal of Cryptology*, vol. 21, no. 2, pp. 149–177, 2008.