Balancing Security and Efficiency for Smart Metering Against Misbehaving Collectors

Jianbing Ni, *Student Member, IEEE*, Kuan Zhang, *Member, IEEE*, Xiaodong Lin, *Fellow, IEEE*, and Xuemin (Sherman) Shen, *Fellow, IEEE*

***Abstract*—Smart grid enables two-way communications between smart meters and operation centers to collect real-time power consumption of customers to improve flexibility, reliability, and efficiency of the power system. It brings serious privacy issues to customers, since the meter readings possibly expose customers' activities in the house. Data encryption can protect the readings, but lengthens the data size. Secure data aggregation improves communication efficiency and preserves customers' privacy, while fails to support dynamic billing, or offer integrity protection against public collectors, which may be hacked in reality. In this paper, we define a new security model to formalize the misbehavior of collectors, in which the misbehaving collectors may launch pollution attacks to corrupt power consumption data. Under this model, we propose a novel privacy-preserving smart metering scheme to prevent pollution attacks for the balance of security and efficiency in smart grid. It achieves end-to-end security, data aggregation, and integrity protection against the misbehaving collectors, which act as local gateways to collect and aggregate usage data and forward to operation centers. As a result, the misbehaving collectors cannot access or corrupt power usage data of customers. In addition, we design a dynamic billing mechanism based on individual power consumption maintained on collectors with the verification of customers. Our analysis shows that the proposed scheme achieves secure smart metering and verifiable dynamic billing against misbehaving collectors with low computational and communication overhead.**

***Index Terms*—Smart grid, smart metering, data aggregation, dynamic billing, security.**

I. Introduction

SMART grid integrates the power grid with information and communication technologies, e.g., network communication, control systems and computation facilities, to achieve two-way electricity and information exchange between operation centers and smart meters, while making the grid more reliable, efficient, secure and greener [1]. It enables operation centers to measure, collect and analyze real-time energy consumption and local electricity generation for energy distribution management, state estimation, outage identification and dynamic billing. The operation centers expose the customers' electricity consumption to power plants, which may help them to adjust the energy production and reduce the need to fire up costly and secondary power plans. The customers not only access their real-time usage information and electricity prices, but also decrease their electricity costs by shifting the uninterrupted activities from peak time to non-peak time.

Although power usage data collection promotes the balance between supply and demand, it brings serious privacy issues toward customers, as it is possible to infer the customers' daily activities, habits and other privacy witnessable references from power consumption data. A relatively low and static daily consumption of a household may indicate that no one is at home [2]; a conspicuous drop of power consumption at midnight may indicate the households go to sleep [3]. The determination of personal behavior patterns is a serious privacy concern in smart grid defined by Electronic Privacy Information Center [4]. To preserve customers' behavior patterns, IEC 62351 [5] resists eavesdropping attacks using TLS encryption [6], including AES CBC, AES GCM or 3DES EDE CBC. Ontario Information Technology Standards [7] suggest to use IPsec or TLS to provide authentication, privacy protection, integrity checking and replay protection for advanced metering communications.

Traditional data encryption increases the data size of consumption reports and causes heavy communication overhead. To address this issue, privacy-preserving data aggregation schemes [8]–[10] have been proposed to compress the consumption reports at local collectors and forward them in a compact form to operation centers. These schemes achieve end-to-end confidentiality of meter readings, but sacrifice the integrity of consumption reports, indicating that they cannot provide sufficient integrity protection on consumption reports against misbehaving collectors. Unfortunately, the consumption reports are transmitted on public networks, such as Cellular and the Internet, with the storage and forwarding of collectors, according to Toronto Hydro. The collectors are vulnerable to be hacked by attackers. Malicious attackers can inject false data into the aggregated reports or corrupt the meter readings without being detected, and thereby affect state estimation, break power dispatch and control electricity prices through misbehaving collectors. The power outage in Ukraine on Dec. 23, 2015 caused by a devastating cyber attack on a power station warns us that any vulnerability in

Manuscript received May 2, 2017; revised July 17, 2017; accepted September 1, 2017. Date of publication October 10, 2017; date of current version February 18, 2019. Paper no. TSG-00586-2017. *(Corresponding author: Xiaodong Lin.)*

J. Ni and X. Shen are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: j25ni@uwaterloo.ca; sshen@uwaterloo.ca).

K. Zhang is with the Department of Electrical and Computer Engineering, University of Nebraska-Lincoln, Omaha, NE 68182 USA (e-mail: kuan.zhang@unl.edu).

X. Lin is with the Department of Physics and Computer Science, Wilfrid Laurier University, Waterloo, ON N2L 3C5, Canada (e-mail: xlin@wlu.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2017.2761804

advanced metering infrastructure may be exploited by hackers to create a blackout. Misbehaving collectors have not being paid enough attentions lately. A handful of schemes [11], [12] aimed to reduce the dependence on a single collector, but bring heavy communication overhead to distribute the reliability to multiple collectors by means of secret sharing. Further, once the consumption reports of different customers are aggregated, it is impossible to achieve dynamic billing for customers. Therefore, it is of importance to design an efficient smart metering scheme that simultaneously supports data aggregation and dynamic billing with high security guarantee.

To balance security and efficiency, we propose a Privacy-Preserving Smart Metering scheme (P^{2}SM) to achieve end-to-end security, data aggregation and dynamic billing, simultaneously. Considering a realistic case that the collectors deployed at public areas may be controlled by attackers, we build a new security model between traditional semi-honest model and malicious model to formally define the misbehavior of collectors. We achieve the authentication, confidentiality and integrity of consumption reports against misbehaving collectors for smart metering based on Chameleon hash function [13], proxy re-encryption [14] and homomorphic authenticators [15]. In addition, we upgrade the collectors with computing and storage resources, such that they can temporarily maintain individual reports for dynamic billing. Our contributions can be summarized as four folds:

- Inspired by the fact that the collectors at public areas may be hacked, we introduce a stronger security model to formalize the collectors' misbehavior in reality. Different from semi-honest adversaries, the misbehaving collectors are not only interested in the personal behavior patterns of customers, but also launch pollution attacks to insert false data into normal meter readings to corrupt state estimation of operation centers.
- To prevent pollution attacks from collectors, we design the P^{2}SM by leveraging proxy re-encryption [14] and homomorphic authenticators [15]. The privacy-preserving data aggregation is achieved to prevent privacy leakage and reduce communication overhead. P^{2}SM does not allow the collectors to generate their signatures by themselves, but aggregate the smart meters' signatures to guarantee the integrity of the aggregated consumption reports. As a result, a misbehaving collector cannot inject false data into the consumption reports or invade the privacy of customers.
- Once smart meters' signatures are aggregated, it is impossible to offer message authentication for customers. We design an identity authentication mechanism from Chameleon hash function [13] for smart metering. With the desirable property of homomorphism, the authentication messages of different customers can be aggregated to further improve communication efficiency.
- To support dynamic billing, P^{2}SM enables collectors to use the maintained individual consumption reports to generate verifiable daily bills for customers. Specifically, the collectors aggregate the consumption reports of each customer with the electricity prices to generate daily bills, and submit them to the operation center. The operation center transforms the encrypted bills to be readable for customers. Furthermore, the customers can verify the correctness of their daily bills to detect the corruption of misbehaving collectors and greedy utilities.

The remainder of the paper is organised as follows. We discuss the related work in Section II, and define system model, security model and design goals in Section III. Then, we describe our P^{2}SM in Section IV and prove the security of P^{2}SM in Section V, followed by the performance evaluation in Section VI. Finally, we draw our conclusion in Section VII.

II. Related Work

Secure communication protocols, such as IPsec or TLS, were suggested by Toronto Hydro [7] to prevent data leakage during transmission, which increase the size of transmitting messages. Privacy-preserving data aggregation schemes [8], [16], [17] have been proposed based on homomorphic encryption to reduce communication overhead and preserve the confidentiality of meter readings in smart grid. Lu *et al.* [8] adopted the Paillier encryption [18] to achieve the aggregation of multi-dimensional data. Fan *et al.* [16] employed the BGN cryptosystem [19] and blinding factors to propose a privacy-enhanced data aggregation scheme to prevent internal attackers from learning the customer's living patterns. Ohara *et al.* [17] adopted the Lifted ElGamal encryption [21] to achieve privacy-preserving data aggregation for meter readings collection. Besides, distributed blinded values are also utilized to support data aggregation and realize efficient communications. Kursawe *et al.* [22] proposed a privacy-preserving data aggregation scheme from blinded values that allows fraud and leakage detection and statistical processing of meter measurements. Lin *et al.* [9] proposed a smart metering system supporting billing and road monitoring using an embedded trusted platform module chip for blinded values generation. Dimitriou and Karame [11] leveraged distributed blinded values to design privacy-preserving aggregation, anonymous tasking and privacy-preserving billing schemes to address the privacy issues in load monitoring and energy trading. Consequently, differential privacy-based data aggregation attracts plenty of attentions and several differentially private data aggregation schemes were proposed to preserve the privacy of customers. Jia *et al.* [23] formalized the human-factor-aware differential aggregation attack in smart grid and designed two privacy-assured aggregation schemes from polynomials and binomial distributed noises to prevent this attack. Won *et al.* [24] designed a differential-private data aggregation protocol for smart metering with fault tolerance by adding distributed Laplace noise. Ni *et al.* [25] proposed a differentially private data aggregation scheme supporting fault tolerance and range-based filtering for advanced metering. Zhang *et al.* [26] introduced a battery-based privacy-preserving scheme to achieve differential privacy and cost saving in smart grid. However, these schemes cannot support dynamic billing since the individual consumption reports are aggregated on the gateways or collectors before being delivered to the operation centers. Although several schemes [9], [11] were designed to support the generation of daily bills from individual measurements, they do not allow the customers to verify whether the bills are correct or not.

To prevent the single point of failure for collectors, Dimitriou and Karame [11] distributed the trust on a single collector to multiple ones by means of secret sharing in smart metering. However, the collectors may be hacked in reality, Rottondi *et al.* [12] proposed an automatic metering scheme by means of verifiable secret sharing to prevent a collusion of collectors from modifying meter readings during smart readings collection, but duplicate collectors are required to be deployed on public areas and the communication burden increases as well. Ni *et al.* [27] introduced the vulnerability of collectors and designed a privacy-enhanced data aggregation scheme from Paillier encryption and homomorphic tags to prevent the false data injection on meter readings from malicious collectors. Jo *et al.* [28] proposed an efficient and privacy-preserving smart metering scheme to resist node compromise attacks and achieve fast authentication for demand response, but this scheme does not support data aggregation.

In addition, the false data injection attack has been widely studied to detect the attackers' misbehavior of manipulation on meter readings at physically protected locations, e.g., smart meters and substations, in electric power grid since introduced by Liu *et al.* [29] in 2009. Liu *et al.* [29] demonstrated the possibility of malicious measurements injection for attackers, and discussed the successful attacks to change the results of state estimation in two plausible attack scenarios, where the attacker can access some specific meters with limited physical protection, and is limited in the resources available to compromise smart meters. To defense this attack, several anomaly detection mechanisms [30], [31] have been proposed based on phaser measurement units (PMU), under the assumption that it is quite difficult to compromise the measurements collected by PMUs. Chen and Abur [32] proposed a PMU placement algorithm, and demonstrated that extra PMUs enable to improve the capability of bad data detection and identification in a power system. Other defense strategies, e.g., spatial-based and temporal-based detection approach [33], generation schedules and synchrophasor data-based detection method [34], and deep learning-based intelligent mechanism [35], were designed to detect the false data injection attacks in power grid. To some extent, our pollution attack is a kind of false data injection attacks, in which the attacker injects false data on the collectors, instead of smart meters. However, privacy-preserving smart metering against false data injection from collectors has not received sufficient attentions currently.

For the above reasons, in this paper, we introduce the security vulnerability of collectors and show a new security model to define the misbehavior of controlled collectors. To resist the pollution attacks, we propose a privacy-preserving smart metering scheme to achieve efficient aggregate authentication for smart meters, and privacy-preserving data aggregation to protect meter readings against misbehaving collectors. Besides, the proposed scheme can support verifiable dynamic billing for customers in smart grid.

III. Problem Statement

In this section, we formalize system model, present security threats and identify design goals.

A. System Model

We formalize power consumption data collection for operation centers and dynamic billing for energy companies as depicted in Fig. 1. Energy companies (utilities) have a good supply of electricity from plants and provide electricity retailing services to customers. To realize real-time power dispatch, operation centers collect, analyze and process real-time power usage data of customers and monitor the power consumption through varying electricity prices. A temper-proof smart meter is installed in each customer's house to measure real-time power consumption and submit the meter readings to operation centers every ρ minutes, in general, $\rho = 15$ or 60. A local collector, which is a wireless access point or base station, is deployed to connect the operation center and smart meters in a home area network. In each reporting time slot, smart meters measure meter readings and deliver consumption reports to the collector through relatively inexpensive WiFi or ZigBee technologies. After receiving the individual consumption reports, the collector transiently stores them, aggregates these reports into a compact one, and delivers the aggregated report to the operation center several times a day through wired network, e.g., the Internet. According to these aggregated reports, the operation center monitors electricity distribution and determines dynamic electricity prices. The electricity price is returned to the collector per day, and the collector computes the daily bills of customers using the maintained individual consumption and sends the electricity bills to the utility. Finally, the customers access their electricity bills via the Internet and regulate their daily activities to decrease electricity costs.

B. Security Model

As the intermediates in advanced metering infrastructure, local collectors are deployed at the public areas and they are vulnerable to be hacked by malicious hackers. The hackers may invade customers' privacy, inject false data, corrupt state estimation and control electricity prices through the misbehaving collectors. To be more close to the reality, we define a new security model with a misbehaving adversary that has rational attack behaviors. The misbehaving collector may be neither completely malicious, to block the power usage data transmission that can be quickly detected by the operation center, nor just honest-but-curious, to be interested in customers' living patterns through eavesdropping. The hacked collector is more powerful than the honest-but-curious adversary, and more rational than the malicious adversary. On one hand, to prevent its misbehavior being identified, a misbehaving collector will follow the communication protocols and pretend to be honest; on the other hand, it tries to use all sorts of methods to achieve the goals of ulterior motives. Specifically, a misbehaving collector launches the following attacks to invade customers' privacy and corrupt state estimation in smart grid:

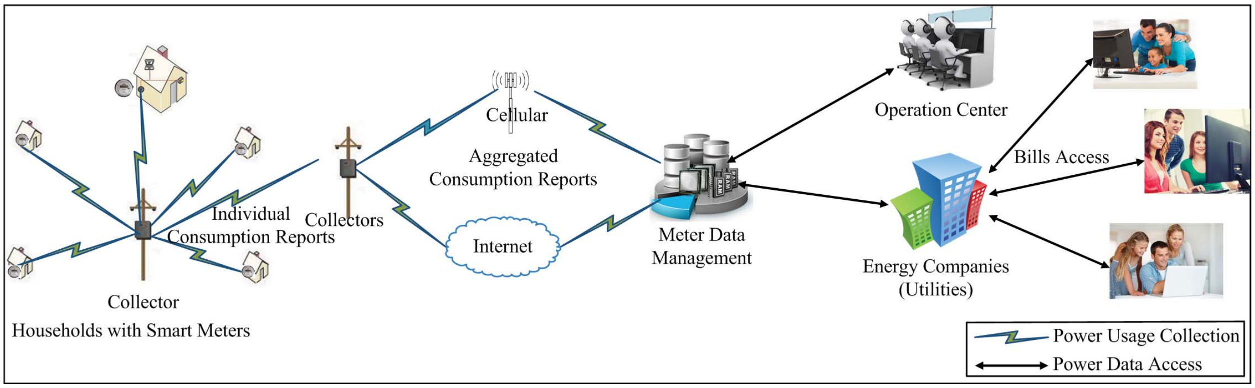


Fig. 1. System Model for Smart Metering.

- A misbehaving collector learns the customers' privacy via eavesdropping.
- A misbehaving collector injects false data into power consumption in home area network to corrupt state estimation or control electricity prices.
- A misbehaving collector may forge the smart meters' individual reports or aggregated reports to corrupt state estimation of the operation center.
- A misbehaving collector may forge the daily electricity bills to cheat the operation center, utilities and customers.

The eavesdropping attack and forgery attack have been discussed in [19] and [20]. The pollution attack is brightly new in our security model. Therefore, we utilize the following game between the misbehaving adversary and the advanced metering infrastructure to formally define this attack:

1) The advanced metering infrastructure setups the whole system to collect the power consumption of customers in a home area network.
2) The adversary can interact with the system and query the individual consumption reports, providing, for each query, a smart meter and its reading. The system generates the individual report for each query and returns the report to the adversary.
3) Finally, the adversary outputs an aggregated report different from the aggregation of queried individual consumption reports.

If the adversary is able to generate a valid aggregated report that is not the aggregation of queried individual reports with non-negligible probability, we say that the adversary wins the game. The smart metering scheme is able to resist pollution attacks if for any misbehaving collector the probability that the collector wins the above game is negligible.

The smart meters are physically protected to prevent customers from stealing electricity. The malfunction of smart meters would be discovered and replaced by utilities in time. In addition, the customers are honest to purchase the electricity from utilities and access their daily electricity bills. The operation center, fully controlled by the government, is honest to manage power transmission and balance the electricity demand and response. The damage of the operation center may directly impact national security and social stability, thereby sufficient security policies are implemented to protect the operation center. The utilities are honest to provide electricity retailing services to customers, while they are curious on customers' privacy and greedy on their benefits, such as increasing their income by modifying customers' bills.

C. Design Goals

To enable privacy-preserving smart metering under the aforementioned system model and resist various security threats, P^{2}SM should achieve the following objectives:

- *Authentication:* To assure that individual consumption reports are from legal customers. It is impossible for an attacker to deliver a forged consumption report acceptable for the operation center.
- *Privacy Preservation:* To guarantee that no attacker is able to learn the meter readings and thereby invade the privacy of customers, even it either eavesdrops on communication channels or hacks the collectors. The curious utility cannot learn the power consumption of their customers, except the daily electricity bills.
- *Integrity Checking:* To ensure that neither individual consumption reports nor aggregated reports can be modified by attackers. Even the misbehaving collectors are not able to corrupt the integrity of consumption reports by injecting false data into normal meter measurements. Therefore, the operation center can obtain correct power consumption data.
- *Dynamic Billing:* To achieve that the daily bills are generated from individual consumption reports and fluctuant electricity prices. The customers can access their bills and verify the correctness.
- *Efficiency:* The communication cost is required to be low to save energy during data transmission and guarantee that the operation center receives the consumption reports within short delay. In addition, there should be no time-consuming operation for smart meters due to their constrained computational capabilities.

IV. The P^{2}SM Scheme

In this section, we describe an overview of P^{2}SM to briefly show the work flow and information flow, and then give the detailed description of P^{2}SM.

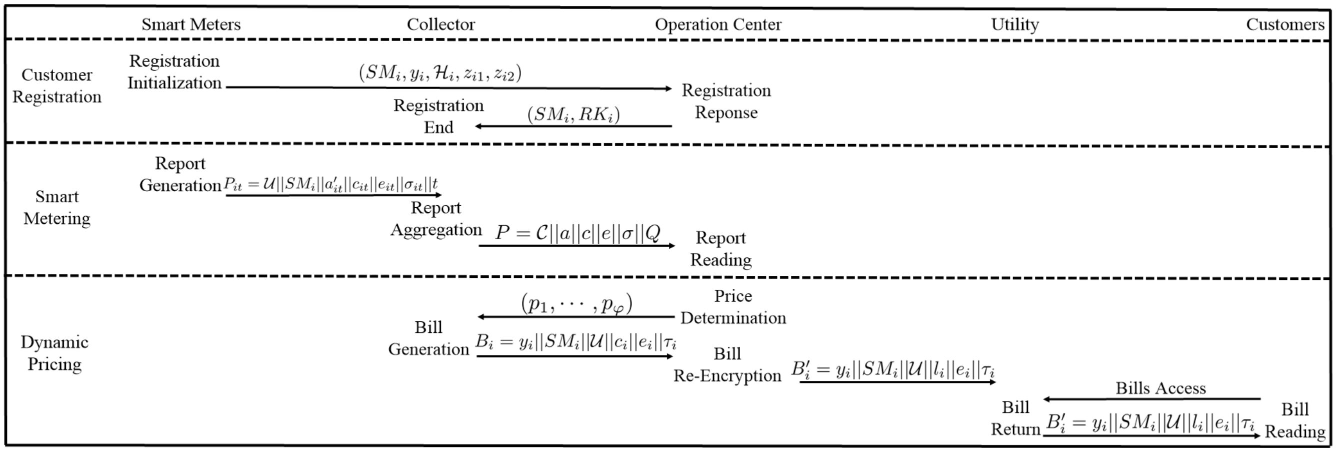

Smart Meters
Collector
Operation Center
Utility
Customers
Customer Registration
Registration Initialization
$(SM_i, y_i, \mathcal{H}_i, z_{i1}, z_{i2})$
Registration Reponse
Registration End
(SM_i, RK_i)
Smart Metering
Report Generation
$P_{it} = \mathcal{U}||SM_i||a'_{it}||c_{it}||e_{it}||\sigma_{it}||t$
Report Aggregation
$P = \mathcal{C}||a||c||e||\sigma||Q$
Report Reading
Dynamic Pricing
$(p_1, \cdots, p_\varphi)$
Price Determination
Bill Generation
$B_i = y_i||SM_i||\mathcal{U}||c_i||e_i||\tau_i$
Bill Re-Encryption
$B'_i = y_i||SM_i||\mathcal{U}||l_i||e_i||\tau_i$
Bills Access
Bill Return
$B'_i = y_i||SM_i||\mathcal{U}||l_i||e_i||\tau_i$
Bill Reading

Fig. 2. Information Flow of P^{2}SM.

A. Overview of P^{2}SM

The reason that the existing privacy-preserving data aggregation schemes in smart grid are vulnerable to the pollution attack is that the collector generates the signature on the aggregated consumption report by itself to ensure the report integrity. If the collector becomes dishonest, it can arbitrarily insert forged data into the aggregated report without being detected by the operation center. To prevent this attack, we extend the homomorphic authenticators in [15] to be pairing-based cryptosystem to achieve the aggregation of the signatures of various measurements generated by different smart meters, which is a big challenge if no parameter is pre-shared among smart meters [36]. To overcome this challenge, we first allow the smart meters to sign the meter measurements rather than their ciphertexts using the individual secret keys based on the homomorphic authenticators [15], and then enable the collectors to re-sign the individual signatures to generate signatures under a common key selected by the operation center for the smart meters in the home area network based on bilinear pairing. Thereby, the re-signed signatures can be aggregated to prevent the misbehaving collector from corrupting the meter measurements. Unfortunately, after the individual signatures are re-signed and aggregated, they cannot offer the authentication for smart meters. To fix this drawback, we design a new identity authentication mechanism based on the Chameleon hash function [13] with batch verification, resulting in the reduction of computational and communication overhead.

In addition, it is of difficulty for the operation center to generate daily bills after the individual consumption reports are compressed. To resolve this problem, we novelly upgrade the capability of collectors with storage spaces. Hence, these individual reports transiently maintained on the collectors can be used to compute the daily bills with the fluctuant electricity prices. To delegate the decryption of daily bills, the proxy re-encryption [14] is leveraged to enable the operation center to re-encrypt the bills generated by the collectors for the customers on behalf of a proxy. With the homomorphism of the proxy re-encryption [14], the ciphertexts of meter readings can be aggregated to improve the communication efficiency. Moreover, to prevent the misbehaving collector from generating cheating bills, the individual signatures of meter measurements are aggregated with the prices to generate verifiable tags on the daily bills. Therefore, the customers can check whether the daily bills are correctly computed and identify the corrupted ones.

The P^{2}SM consists of six phases, namely, System Initialization, Customer Registration, Report Generation, Report Aggregation, Report Reading and Dynamic Billing. The information flow of P^{2}SM is depicted in Fig. 2.

- System Initialization: The operation center bootstraps the whole system for smart metering and generates the public parameters *Params* and its secret-public key pair (k, K).
- Customer Registration: The customer C_i with an installed smart meter SM_i on the house registers at the operation center using the registration message $(SM_i, y_i, \mathcal{H}_i, z_{i1}, z_{i2})$, in which $\mathcal{H}_i$ is the commitment generated from the Chameleon hash function and (z_{i1}, z_{i2}) is the ciphertext of a random key k. The operation center returns (SM_i, RK_i) to the local collector, where RK_i is the re-sign key used to transform C_i's signature to a signature under the common key α selected by the operation center for the smart meters in the home area network.
- Report Generation: The smart meter SM_i reads the measurement m_{it} at a time slot t and generates a consumption report $P_{it} = \mathcal{U}||SM_i||a'_{it}||c_{it}||e_{it}||\sigma_{it}||t$, in which a'_{it} is the authentication message, (c_{it}, e_{it}) is the ciphertext of m_{it} and σ_{it} is the signature on m_{it}. SM_i sends P_{it} to the local collector.
- Report Aggregation: The collector aggregates the authentication messages, ciphertexts and signatures in the individual consumption reports during each forwarding period Q to generate an aggregated report $P = \mathcal{C}||a||c||e||\sigma||Q$ using the re-sign keys RK_i of all smart meters in its home area network, and forwards P to the operation center.
- Report Reading: The operation center checks the validity of the aggregated authentication message a, decrypts the aggregated ciphertext (c, e) and verifies

the aggregated signature σ. Finally, the operation center obtains the total power consumption m in the home area network for state estimation and demand response.

- Dynamic Billing: The operation center determines the fluctuant electricity prices $(p_1, \ldots, p_\varphi)$ during a day. The collector aggregates the ciphertexts of meter readings with the prices to generate the daily bill (c_i, e_i) for the customer, and aggregates the signatures with the prices to obtain a verifiable tag τ_i on the daily bill. To enable the customer to read the bill, the operation center re-encrypts the daily bill $B_i = y_i||SM_i||\mathcal{U}||c_i||e_i||\tau_i$ to generate $B_i' = y_i||SM_i||\mathcal{U}||l_i||e_i||\tau_i$ for the customer. Therefore, the customer can read the daily bill and uses τ_i to check the correctness of the daily bill.

B. The Detailed P²SM

We then describe the construction of P^{2}SM in detail.

1) System Initialization: The operation center (OC) provides electricity distribution and demand response for the customers $\mathbb{C} = \{C_1, \ldots, C_N\}$ in the residential area $\mathbb{RA}$. Suppose that these customers buy electricity from a utility $\mathcal{U}$, (it is compatible that $\mathbb{C}$ use the power offered by multiple utilities). OC bootstraps the advanced metering infrastructure on behalf of a trust authority. Concretely, OC first determines the security parameter κ, which denotes the security level of the system and κ is 160 or 256 usually. OC chooses a large prime p, where $|p| = \kappa$. OC also generates two cyclic groups $(\mathbb{G}, \mathbb{G}_T)$ with the same order p. g, g_0 are two generators of $\mathbb{G}$, and $\hat{e} : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$ is a bilinear map. $H : \{0,1\}^* \to \mathbb{G}$ and $H_1 : \{0,1\}^* \to \{0,1\}^\kappa$ are cryptographic hash functions and $F : \mathbb{G} \times \{0,1\}^* \to \mathbb{Z}_p$ is a pseudo-random function. (E_s, D_s) are the encryption and decryption algorithms of AES cryptosystem. Then, OC randomly chooses $k \in \mathbb{Z}_p$ to compute $K = g^k$. Finally, OC releases the public parameters:

$$Params = \{p, \mathbb{G}, \mathbb{G}_T, \hat{e}, g, g_0, H, H_1, F, E_s, D_s, K\},$$

and keeps the secret key k in private.

2) Customer Registration: When a customer $C_i \in \mathbb{C}$'s house in the $\mathbb{RA}$ connects the smart grid, OC installs a smart meter SM_i for C_i. In the registration, C_i first randomly chooses $x_i \in \mathbb{Z}_p$ as the private key and computes the corresponding public key as $y_i = g^{x_i} \in \mathbb{G}$. The secret key x_i is plugged into SM_i or stored in a trusted platform module (TPM) integrated into the smart meter SM_i and the public key certificate is publicly accessed by any entity. Then, SM_i randomly picks $a_i, b_i \in \mathbb{Z}_p$ to compute a Chameleon hash value as $\mathcal{H}_i = g^{a_i} y_i^{b_i}$. After that, SM_i chooses two random values $k_i, r_i \in \mathbb{Z}_p$ to calculate $z_{i1} = g^{r_i}$, $r_i' = H_1(z_{i1}, K^{r_i})$ and $z_{i2} = E_s(r_i', k_i)$. Finally, SM_i sends $(SM_i, y_i, \mathcal{H}_i, z_{i1}, z_{i2})$ to OC, and keeps (a_i, b_i, k_i) in the TPM, along with x_i.

Upon receiving $(SM_i, y_i, \mathcal{H}_i, z_{i1}, z_{i2})$, OC decrypts (z_{i1}, z_{i2}) to obtain the tag k_i as $r_i' = H_1(z_{i1}, z_{i1}^k)$ and $k_i = D_s(r_i', z_{i2})$. Then, OC randomly picks $\alpha \in \mathbb{Z}_p$ as a unique identifier of $\mathbb{RA}$ to compute a re-sign key $RK_i = y_i^\alpha$, if C_i is the first customer in $\mathbb{RA}$; otherwise, U_i uses the existing α to compute RK_i. Finally, OC sends (SM_i, RK_i) to the local collector in $\mathbb{RA}$ through a secure channel, and stores $(SM_i, y_i, \mathcal{H}_i)$ in its database and keeps $(T_i, RK_i, \alpha, g^\alpha)$ secretly.

3) Report Generation: To achieve real-time power dispatch, smart meters measure power consumption and deliver electricity consumption reports every ρ minutes, i.e., $\rho = 15$ or 60 ($\rho = 60$ for Toronto Hydro [7]). Suppose that a smart meter SM_i measures the meter reading m_{it} at a time slot t. SM_i generates an individual consumption report as follows:

- Use $b_{it}' = F(H(SM_i||k_i), t)$ to compute the authentication message $a_{it}' = x_i \cdot (b_i - b_{it}') + a_i \mod p$;
- Randomly pick $s_{it} \in \mathbb{Z}_p$ to generate the ciphertext of meter reading m_{it} as:

$$c_{it} = K^{s_{it}}, \quad e_{it} = \hat{e}\big(g_0^{m_{it}} g^{s_{it}}, g\big);$$

- Use x_i to generate a signature on the meter reading as:

$$\sigma_{it} = \Big(H(SM_i||\mathcal{U}||t) g_0^{a_{it}'} g^{m_{it}}\Big)^{\frac{1}{x_i}}; \qquad (1)$$

- Send the individual consumption report $P_{it} = \mathcal{U}||SM_i||a_{it}'||c_{it}||e_{it}||\sigma_{it}||t$ to the collector in this area.

4) Report Aggregation: The collector transiently maintains the received individual consumption reports from smart meters. It is required to forward the consumption reports to OC φ times per day ($\varphi = 5$ or 24). In each forwarding period Q, the collector aggregates the individual consumption reports received in Q from N smart meters in $\mathbb{RA}$ into an aggregated report P as follows:

$$c = \prod_{t\in Q}\prod_{i=1}^{N} c_{it}; \quad e = \prod_{t\in Q}\prod_{i=1}^{N} e_{it}; \qquad (2)$$

$$a = \sum_{t\in Q}\sum_{i=1}^{N} a_{it}' \mod p; \quad \sigma = \prod_{t\in Q}\prod_{i=1}^{N} \hat{e}(\sigma_{it}, RK_i). \qquad (3)$$

The collector sets $P = \mathcal{C}||a||c||e||\sigma||Q$ and forwards P to OC, where $\mathcal{C}$ is the identifier of the collector.

5) Report Reading: After receiving $P = \mathcal{C}||a||c||e||\sigma||Q$, OC performs the following steps to read the aggregated report P:

- Use each customer's unique tag k_i to compute $b_{it}^* = F(H(SM_i||k_i), t)$ and verify whether all reports are released by legitimate smart meters by checking the equation (4):

$$\prod_{t\in Q}\prod_{i=1}^{N} \mathcal{H}_i \overset{?}{=} g^a \cdot \prod_{t\in Q}\prod_{i=1}^{N} y_i^{b_{it}^*}. \qquad (4)$$

  If the equation (4) holds, continue to decrypt (c, e); otherwise, retrieve the individual consumption reports from the collector to find the invalid reports.
- Decrypt the aggregated ciphertext (c, e) as $M = e\hat{e}(c, g)^{-\frac{1}{k}}$ and recover the discrete log of M base $\hat{e}(g_0, g)$ using Pollard's lambda method [37] to obtain $m = \sum_{t\in Q}\sum_{i=1}^{N} m_{it}$.
- Verify whether the equation (5) is valid or not:

$$\sigma \overset{?}{=} \hat{e}\left(\prod_{t\in Q}\prod_{i=1}^{N} H(SM_i||\mathcal{U}||t) g_0^a g^m, g^\alpha\right). \qquad (5)$$

If yes, accept m, which is the total power consumption of customers in $\mathbb{RA}$ in the period Q; otherwise, reject m and retrieve the individual consumption reports from the collector to find the corrupted reports utilizing a recursive divide-and-conquer approach.

6) Dynamic Billing: According to the power consumption of customers in $\mathbb{RA}$, OC determines the electricity price in every forwarding period during a day, that is, $(p_1, \ldots, p_\varphi)$, where p_j denotes the electricity price in the jth forwarding period Q_j, and sends $(p_1, \ldots, p_\varphi)$ to the collector. The collector aggregates the individual consumption reports of a customer with the electricity prices to generate an electricity bill for the customer. Specifically, for a customer C_i, the collector computes

$$c_i = \prod_{j=1}^{\varphi} \prod_{t \in Q_j} c_{it}^{p_j}, \;\; e_i = \prod_{j=1}^{\varphi} \prod_{t \in Q_j} e_{it}^{p_j}, \;\; \tau_i = \prod_{j=1}^{\varphi} \prod_{t \in Q_j} \sigma_i^{p_j}, \tag{6}$$

where $t \in Q_j$ means that the time slot t is in the reporting period Q_j, and delivers the bill $B_i = y_i||SM_i||\mathcal{U}||c_i||e_i||\tau_i$ to OC. Then, OC verifies the correctness of the bills in $\mathbb{RA}$ by verifying the equation (7):

$$\hat{e}(g_0, g)^{\sum_{j=1}^{\varphi} m_j p_j} = \prod_{i=1}^{N} e_i \hat{e}(c_i, g)^{-\frac{1}{k}}, \tag{7}$$

where m_j is the total power consumption of customers in $\mathbb{RA}$ in the period Q_j. If it holds, OC further computes $l_i = \hat{e}(c_i, y_i)^{\frac{1}{k}}$ and sends the bill $B_i' = y_i||SM_i||\mathcal{U}||l_i||e_i||\tau_i$ to $\mathcal{U}$. In addition, OC can delegate $\mathcal{U}$ to perform proxy re-encryption to transform the ciphertexts of OC to be decryptable for C_i on behalf of a proxy. Specifically, OC sends $USK_i = y_i^{\frac{1}{k}}$ to $\mathcal{U}$ to enable it to compute $l_i = \hat{e}(c_i, USK_i)$ for C_i. Finally, C_i decrypts (l_i, e_i) by computing $D_i = e_i l_i^{-\frac{1}{x_i}}$ and recovering the discrete log of D_i base $\hat{e}(g_0, g)$ using Pollard's lambda method [37] to obtain $d_i = \sum_{j=1}^{\varphi} \sum_{t \in Q_j} m_{it} p_j$. To verify the correctness of d_i, C_i checks the equation (8):

$$\hat{e}(\tau_i, y_i) \stackrel{?}{=} \hat{e}\left(\prod_{j=1}^{\varphi} \prod_{t \in Q_j} H(SM_i||\mathcal{U}||t)^{p_j} g_0^{\sum_{j=1}^{\varphi} \sum_{t \in Q_j} a_{it}' p_j} g^{d_i}, g \right), \tag{8}$$

where $a_{it}' = x_i(b_i - F(H(SM_i||k_i), t) + a_i) \bmod p$. If the equation (8) holds, C_i accepts the bill d_i; otherwise, rejects it.

V. Security Analysis

In this section, we analyze the security properties of P^{2}SM, following the security goals described in Section III, including authentication, confidentiality and integrity.

- *Authentication:* In customer registration, SM_i utilizes the ElGamal encryption to send k_i to OC. Since the ElGamal encryption is semantically secure against chosen plaintext attacks [21] based on Hash-Diffie-Hellman problem, only OC is able to recover k_i. Hence, k_i is shared between SM_i and OC. To achieve efficient authentication, the Chameleon hash function is leveraged to design the interactions between SM_i and OC. Firstly, $\mathcal{H}_i = g^{a_i} y_i^{b_i}$ is one-way, indicating that it is easy to compute $\mathcal{H}_i$ from (a_i, b_i), but no one can recover (a_i, b_i) from $\mathcal{H}_i$, as long as the Discrete Logarithm (DL) assumption [13] holds. In addition, no one is able to find a collision (a_i', b_i') of (a_i, b_i) to make $\mathcal{H}_i = g^{a_i'} h_i^{b_i'}$ hold without x_i in polynomial time with non-negligible probability, unless the DL problem is tractable. However, having x_i, SM_i can compute a_i from any given b_i. Therefore, if the ElGamal encryption is semantically secure and the DL problem is intractable, it is impossible for an adversary to pretend a legal smart meter to generate consumption reports without being detected by OC.
- *Confidentiality:* To prevent the customers' power consumption from being revealed, we adopt to the proxy re-encryption [14] to encrypt meter readings. Since the proxy re-encryption is proved secure against chosen plaintext attacks, the confidentiality of m_{it} is satisfied to prevent attackers from invading C_i's privacy, even attackers can eavesdrop on communication channels and capture the ciphertexts. When the collector obtains all individual consumption reports from smart meters in $\mathbb{RA}$, it cannot recover the meter readings but aggregating the ciphertexts based on additive homomorphism to reduce communication overhead. As for OC, it can recover the sum of power consumption in $\mathbb{RA}$ by using its secret key. In Dynamic Billing phase, the collector aggregates the individual consumption reports with the electricity prices to generate the bills and forwards the results to OC or U to allow them to re-encrypt the bills to be decryptable for the customers. As the proxy re-encryption is secure based on the Computational Bilinear Inverse Diffie-Hellman (BIDH) assumption [14], the meter readings and the electricity bills are confidential against eavesdroppers and curious entities.
- *Integrity:* To resist pollution attacks, P^{2}SM should ensure the integrity of consumption reports from smart meters to the operation center. The signatures of smart meters are used to ensure the integrity of individual consumption reports, and the aggregated signature is generated from the signatures of smart meters by the collector to prevent data corruption during the transmission from the collector to the operation center. Thus, the integrity of reports depends on the unforgeability of both the individual signatures and the aggregated signature. We prove the unforgeability of the individual signatures and the aggregated signature separately.

To ensure the integrity of the individual consumption report, SM_i generates a signature using its private key as $\sigma_{it} = (H(SM_i||\mathcal{U}||t) g_0^{a_{it}'} g^{m_{it}})^{\frac{1}{x_i}}$. The unforgeability of this signature can be reduced to the Diffie-Hellman Inversion (DHI) assumption [38], that is, within non-negligible advantage, there is no probabilistic polynomial-time algorithm to solve DHI problem: given $h, h^s \in \mathbb{G}$, where $s \in \mathbb{Z}_p$, to compute $h^{\frac{1}{s}} \in \mathbb{G}$.

Theorem 1: The signature in the individual consumption report is existentially unforgeable against adaptive chosen message attacks under the security model [20], provided that the DHI problem is difficult to be addressed with a non-negligible probability in probabilistic polynomial time.

Proof: Suppose that an adversary $\mathcal{A}$ can break the existential unforgeability of the signature with a non-negligible probability, then we can construct an algorithm $\mathcal{B}$ to solve the DHI problem. Let h be a generator of $\mathbb{G}$. $\mathcal{B}$ is given $h, h^s \in \mathbb{G}$, where $s \in \mathbb{Z}_p$, its goal is to compute $h^{\frac{1}{s}}$. $\mathcal{B}$ simulates a challenger and interacts with $\mathcal{A}$ in the following way.

- In setup, $\mathcal{B}$ sets the public key v to $h^{\frac{s}{r}}$ and the parameters g to h^{sr_1}, g_0 to h^{sr_2}, where r, r_1, r_2 are random values chosen from $\mathbb{Z}_p$, and forwards them to $\mathcal{A}$.
- $\mathcal{B}$ programs a random Oracle to answer hash queries. To ensure the consistency, it maintains a list of tuples to keep the queries and corresponding responses. When receiving queries $(SM_i, \mathcal{U}, t)$ from $\mathcal{A}$, $\mathcal{B}$ flips a bias coin $\theta_i \in \{0, 1\}$, such that $\Pr[\theta_i = 0] = 1/(q_s + 1)$, where q_s is the maximum of signing queries that $\mathcal{A}$ can make. If $\theta_i = 0$, $\mathcal{B}$ computes $w_i = h^{\beta_i}$; otherwise, $\theta_i = 1$ and $\mathcal{B}$ computes $w_i = h^{s\beta_i}$, where β_i is a random value chosen from $\mathbb{Z}_p$. At last, $\mathcal{B}$ adds a tuple $(SM_i, \mathcal{U}, t, \theta_i, \beta_i, w_i)$ to the list, and returns w_i to $\mathcal{A}$.
- $\mathcal{B}$ also programs a signing Oracle and maintains a list of tuples to keep the queries and responses. When $\mathcal{A}$ queries $(SM_i, \mathcal{U}, t, a'_{it}, m_{it})$, $\mathcal{B}$ firstly checks the list in hash queries. If $(SM_i, \mathcal{U}, t)$ has not been queried, $\mathcal{B}$ generates the corresponding (θ_i, β_i, w_i) for $(SM_i, \mathcal{U}, t)$. If $\theta_i = 0$, $\mathcal{B}$ aborts and returns failure; If $\theta_i = 1$, $\mathcal{B}$ sets $\sigma_{it} = h^{\beta_i r + r_1 a'_{it} r + r_2 m_{it} r}$. Observe that σ_{it} is a valid signature on $(SM_i, \mathcal{U}, t, a'_{it}, m_{it})$ under the public key $h^{\frac{s}{r}}$. Finally, $\mathcal{B}$ returns σ_{it} and adds $(SM_i, \mathcal{U}, t, \theta_i, \beta_i, a'_{it}, m_{it}, \sigma_{it})$ to the list.
- Eventually, $\mathcal{A}$ produces a message-signature pair $(SM_i, \mathcal{U}, \hat{t}, \hat{a}'_i, \hat{m}_i, \hat{\sigma}_i)$, such that no signature query has been made for $(SM_i, \mathcal{U}, \hat{t}, \hat{a}'_i, \hat{m}_i)$. If there is no tuple in hash list, $\mathcal{B}$ issues $(SM_i, \mathcal{U}, \hat{t})$ to hash query. $\mathcal{B}$ aborts and returns failure, if $\hat{\sigma}_i$ is invalid. Next, $\mathcal{B}$ finds the tuple on hash list. If $\hat{\theta}_i = 1$, $\mathcal{B}$ aborts and returns failure; otherwise, $\hat{\theta}_i = 0$ and therefore $H(SM_i||\mathcal{U}||\hat{t}) = h^{\hat{\beta}_i}$. Hence, $\hat{\sigma}_i = h^{\frac{\hat{\beta}_i r}{s}} h^{r_1 \hat{a}'_i r + r_2 \hat{m}_i r}$. Then, $\mathcal{B}$ outputs the required $h^{\frac{1}{s}} = (\hat{\sigma}_i h^{-(r_1 \hat{a}'_i r + r_2 \hat{m}_i r)})^{\frac{1}{\hat{\beta}_i r}}$.

Therefore, if the DHI problem cannot be solved with a non-negligible probability in probabilistic polynomial time, no adversary can forge the signatures on individual reports. ■

The integrity of bills can be reduced to the DHI assumption as the signatures on bills are the aggregation of smart meter's signatures. If the single signature σ_{it} is unforgeable, its aggregated signature τ_i cannot be forged by attackers as well.

Then, we show that it is impossible to forge a valid aggregated report-signature pair (σ, m) in probabilistic polynomial time under the assumption of Conference-Key Sharing (CONF) [39] in group $\mathbb{G}_T$, that is, there is no probabilistic polynomial-time algorithm that solves CONF problem [39] within a non-negligible probability: given $g, g^s g^{sv} \in \mathbb{G}$, where $s, v \in \mathbb{Z}_p$, to compute $\hat{e}(g, g)^v \in \mathbb{G}_T$.

Theorem 2: The probability of generating a valid aggregated signature σ, which is not equal to the aggregation of smart meters' signatures, in probabilistic polynomial time is negligible, provided that the CONF problem is hard.

Proof: If there is a probabilistic polynomial-time adversary $\mathcal{A}$ can break the unforgeability of the aggregated signature within a non-negligible probability, we can construct an algorithm $\mathcal{B}$ to solve the CONF problem.

Let g be a generator of $\mathbb{G}$. $\mathcal{B}$ is given $g, D = g^s, D_1 = g^{sv} \in \mathbb{G}$, where $s, v \in \mathbb{Z}_p$, its goal is to compute $D_2 = \hat{e}(g, g)^v$. $\mathcal{B}$ simulates a challenger, who is allowed to access the signing Oracle $\mathcal{SO}$ that can output the signatures on individual reports, and interacts with the adversary $\mathcal{A}$ as follows.

- In setup, $\mathcal{B}$ randomly chooses $r_i \in \mathbb{Z}_p$ to set the public key h_i to D^{r_i} and the re-sign key RK_i to $D_1^{r_i}$, for $1 \le i \le N$. Then, $\mathcal{B}$ randomly picks $\gamma \in \mathbb{Z}_p$ to set the parameter g_0 to g^γ. Finally, $\mathcal{B}$ sends $(\{h_i, RK_i\}_{1 \le i \le N}, g, g_0)$ to $\mathcal{A}$.
- $\mathcal{A}$ queries the signatures on individual reports under any public key in $\{h_i\}$ for $1 \le i \le N$. $\mathcal{B}$ issues a signature query to $\mathcal{SO}$ and receives σ_i, then, returns σ_i to $\mathcal{A}$.
- Eventually, $\mathcal{A}$ produces an aggregated signature $\hat{\sigma}$ on the compressed reports $(SM_i, \mathcal{U}, t, \hat{a}, \hat{m})$ for $1 \le i \le N$ in a time period $t \in Q$. Suppose that $\hat{\sigma}$ is a valid signature on $\hat{m}$; otherwise, $\mathcal{B}$ reports failure and aborts. Thus, $\hat{\sigma}$ satisfies the verification equation, i.e.,
$$\hat{\sigma} = \hat{e}\left(\prod_{t \in Q} \prod_{i=1}^{N} H(SM_i||\mathcal{U}||t) g^{\hat{a}} g_0^{\hat{m}}, g^v\right).$$
Assume the expected signature, which would be obtained from the honest smart meters, be σ on the report $(SM_i, \mathcal{U}, t, a, m)$ for $1 \le i \le N$ in a time period $t \in Q$. σ also satisfies the verification equation, i.e.,
$$\sigma = \hat{e}\left(\prod_{t \in Q} \prod_{i=1}^{N} H(SM_i||\mathcal{U}||t) g^{a} g_0^{m}, g^v\right).$$
If $\hat{a} = a$ and $\hat{m} = m$, then $\hat{\sigma} = \sigma$. Define $\Delta a = \hat{a} - a$ and $\Delta m = \hat{m} - m$, then, either Δa or Δm is nonzero.
  - If $\hat{\sigma} \neq \sigma$, we divide the verification equation for $\hat{\sigma}$ by the equation for σ and obtain
$$\hat{\sigma}/\sigma = \hat{e}\left(\prod_{t \in Q} \prod_{i=1}^{N} g^{\Delta a} g_0^{\Delta m}, g^v\right).$$
Since $g_0 = g^\gamma$, we have
$$\hat{\sigma}/\sigma = \hat{e}\left(g^{\sum_{t \in Q} \sum_{i=1}^{N} \Delta a + \gamma \Delta m}, g^v\right).$$
Rearranging the equation yields
$$D_2 = \hat{e}(g, g)^v = \left(\frac{\sigma}{\hat{\sigma}}\right)^{\sum_{t \in Q} \sum_{i=1}^{N} \Delta a + \gamma \Delta m},$$
which is the solution to the CONF problem.
  - Otherwise, we get $\hat{e}(g^{\sum_{t \in Q} \sum_{i=1}^{N} \Delta a + \gamma \Delta m}, g^v) = 1$ and
$$D_2 = \hat{e}(g, g)^v = 1^{\sum_{t \in Q} \sum_{i=1}^{N} \Delta a + \gamma \Delta m}.$$
So we can solve the CONF problem.

Therefore, if the CONF problem cannot be solved within a non-negligible probability in probabilistic polynomial time, any adversary cannot corrupt the aggregated reports. ■

TABLE I
COMPARISON OF TIME COSTS

Unit: ms, $N = 100$

Phase	Report Generation	Report Aggregation	Report Reading	Dynamic Billing		
	Smart Meter	Collector	Operation Center	Collector	Operation Center	Customer
P^2SM	24.7	2355.4	328.1	2254.7	3564.1	54.9
EPPA [8]	26.2	4906.3	97.7	Null	Null	Null
Fan14 [16]	7.3	6737.6	130.9	Null	Null	Null
Ohara14 [17]	96.2	25.4	174.3	Null	26.5	183.4

In summary, P^2SM can achieve the authentication, confidentiality and integrity of the individual consumption reports and aggregated reports, as well as the electricity bills. Thus, the misbehaving collector cannot corrupt both consumption reports and electricity bills without being detected.

VI. PERFORMANCE EVALUATION

In this section, we evaluate the performance of P^2SM in terms of computational, communication and storage burden, and discuss the implementation on the current advanced metering infrastructure.

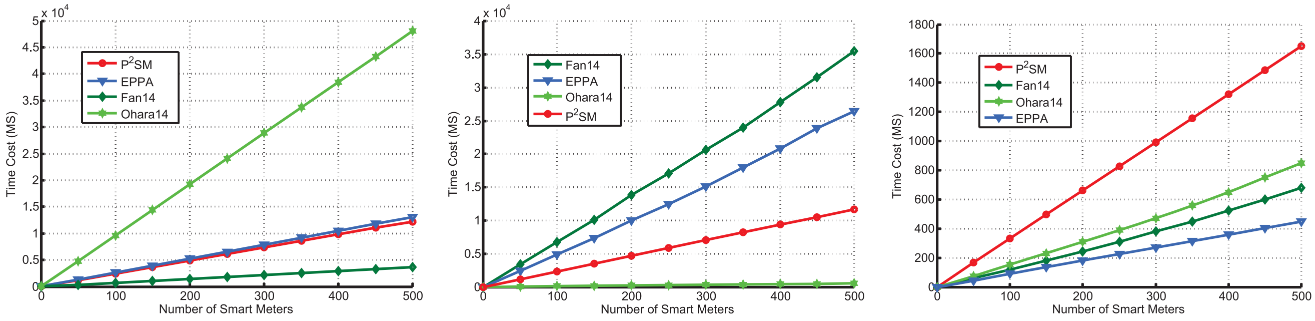
A. Computational Cost

To evaluate the computational cost of P^2SM, we count the number of time-consuming operations on elliptic curve groups, including scalar multiplication (SM), point addition (PA), hash to point (HP), bilinear pairing (BP) and multiplication in $\mathbb{G}_T$ (MU_T). When a customer C_i registers on OC, it is required to perform $3SM + HP + 2PA$ operations to generate $(\mathcal{H}_i, z_{i1}, z_{i2})$, and OC runs $2SM + PA$ operations to recover k_i and compute RK_i. In each reporting slot t, SM_i computes $(a'_{it}, c_{it}, t_{it}, \sigma_{it})$ to generate P_{it}, which needs SM_i to execute $4SM + HP + 2PA + 2MU_T$ operations. Here $\hat{e}(g_0, g)$ and $\hat{e}(g, g)$ can be pre-computed in System Initialization phase to reduce the computational overhead of SM_i. To aggregate the individual consumption reports, the collector performs $(2N|Q| - 2)PA + N|Q|BP$ operations to generate P, where $|Q|$ denotes the number of individual reports of SM_i received in a forwarding period Q. Finally, OC executes $(2N|Q| + 1)PA + (N|Q| + 4)SM + N|Q|HP + MU_T + BP$ and discrete logarithm operations to obtain the sum of power consumption in $\mathbb{RA}$, if all the reports are valid. Otherwise, OC needs to execute $N|Q|(2SM + PA)$ operations to find the invalid authentication messages if the equation (4) does not hold; or $N|Q|(2PA + 2SM + HP + 2BP)$ operations to identify the invalid signatures if the equation (5) does not hold. In Dynamic Billing phase, for a customer C_i, the collector needs to aggregate the power consumption by performing $72|Q_j|SM + (24|Q_j| - 3)PA$ operations, where $|Q_j|$ denotes the number of reporting slots in each forwarding slot. Then, OC runs $NSM + (2N - 1)MU_T$ operations to verify the correctness of electricity bills and executes $SM + BP$ operations to generate l_i or $\mathcal{U}$ helps OC to compute BP operation. Finally, C_i performs MU_T and discrete logarithm operations to recover the bill d_i, and checks the correctness of d_i by running $(24|j| + 2)SM + 2PA + 24|j|HP + 2BP$ operations.

We conduct an experience on a notebook with Intel Core i5-4200U CPU @ 2.29GHz and 4.00GB memory. We use the MIRACL library to implement number-theoretic based methods of cryptography. The Weil pairing is utilized to realize the bilinear pairing and the elliptic curve is chosen with a base field size of 512 bits. The parameter p is 160 bits. We compare the P^2SM with three schemes, EPPA [8] (based on Paillier encryption [18]), Fan14 [16] (based on BGN encryption [19]) and Ohara14 [17] (based on Lifted ElGamal encryption [21]). While P^2SM is constructed from proxy re-encryption [14]. To keep the consistency, we utilize the same settings of these schemes in the experience. The number of customers in $\mathbb{RA}$ is 100, the number of reporting slots in a period Q is 1 and the number of reporting period per day φ is 24. The execution time of each entity in report generation, report aggregation, report reading and dynamic billing phases of four schemes are shown in Table I. Fig. 3 shows the comparison results of four schemes about the computational cost for each entity. Although Fan14 [16] is the most efficient scheme in report generation, since it utilizes the BGN encryption to encrypt the meter measurements and no commitment is generated, it is the most inefficient one in report aggregation. EPPA [8] is the fastest scheme in report reading, as no discrete logarithm computation is needed and less bilinear pairings are computed compared with P^2SM, but it is time-consuming in report aggregation. Moreover, Fan14 [16] and EPPA [8] do not achieve dynamic billing for customers. Ohara14 [17] has a good performance on the computational overhead, but it is vulnerable to the pollution attacks from misbehaving collectors, as well as Fan14 [16] and EPPA [8]. P^2SM is not the most efficient one in four schemes, even the least efficient one in report reading, because it utilizes the pairing-based cryptosystem to offer a higher security guarantee compared with Fan14 [16], EPPA [8] and Ohara14 [17]. Moreover, the bottleneck of the smart metering on computational capability is the smart meters, while our scheme is very efficient on report generation for smart meters. P^2SM is still efficient since the time cost on report reading is only 328ms, while the operation center is always powerful on computation. The collector's computational overhead can be reduced by decreasing the number of smart meters in its coverage area and building the hierarchical network structure to improve the efficiency of meter measurement collection.

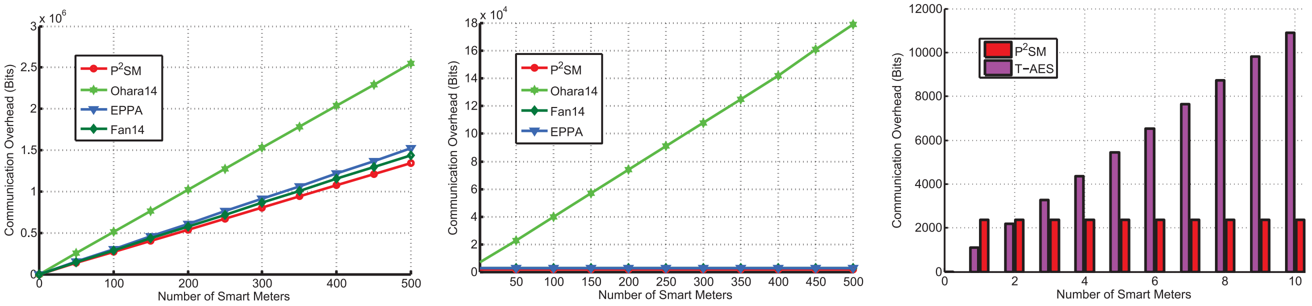
B. Communication and Storage Overhead

The communications of P^2SM composes of SM-to-Collector communication and Collector-to-OC communication. In the SM-to-Collector communication, SM_i is required to send 2688-bit P_{it} at a time slot t to the collector, where $\mathcal{U}$, SM_i and t are assumed to be 160 bits, respectively. In the



(a) Computational Cost of Smart Meters (b) Computational Cost of the Collector (c) Computational Cost of the Operation Center

Fig. 3. Comparison on Computational Overhead.



(a) Overhead between SM and Collector (b) Overhead between Collector and OC (c) Comparison of P^{2}SM and TLS protocol

Fig. 4. Comparison on Communication Overhead.

Collector-to-OC communication, all P_{it} are aggregated to be P, which is only 2388 bits, if $\mathcal{C}$ and Q are 160 bits, respectively. Therefore, the communication overhead is significantly reduced through data aggregation. In addition, the collector aggregates the individual reports to generate a bill for each customer every day. The bill is only 2880 bits. Fig. 4 shows the comparison results of four schemes about the communication overhead of SM-to-Collector communication (Fig. 4(a)) and Collector-to-OC communication (Fig. 4(b)). Since the ciphertext of proxy re-encryption [14] is shorter than those of Paillier encryption [18], BGN encryption [19] and Lifted ElGamal encryption [21] (the commitment in [21] is 1024 bits), the communication burden of SM-to-Collector communication in P^{2}SM is lower than those in the other three schemes as shown in Fig. 4(a). After the individual consumption reports are aggregated, the communication overhead becomes constant, which is still lower than those in EPPA [8], Fan14 [16] and Ohara14 [17] (the overhead of Collector-to-OC communication is linear to the number of smart meters in the home area network). Fig. 4(c) shows the comparison on communication overhead of P^{2}SM and TLS protocol (T-AES) [6], in which AES-256 is used to encrypt meter readings and BLS signature [20] is used to guarantee authentication and data integrity. If $N > 2$ in $\mathbb{RA}$, P^{2}SM is more efficient than TLS protocol on Collector-to-OC communication.

In addition, the collector needs sufficient storage space to transiently maintain the individual reports in $\mathbb{RA}$. If $N = 1000$ in $\mathbb{RA}$ and $\rho = 15$, these individual reports would possess 30.8MB storage space every day. Therefore, each collector only needs to deploy 61.6MB memory to support power consumption collection and dynamic billing for customers.

C. Implementation

We give a detailed description to show how to implement our P^{2}SM. The public-key infrastructure (PKI) is deployed to facilitate secure information transmission. In PKI, a Certificate Authority (CA) issues public certificates to smart meters and the operation center. All the secret information (e.g., secret keys and random values) are maintained on secure chips (e.g., TMP). The TMP can be used to perform the cryptographic operations. The security parameter κ is 160 or 256. The elliptic curve can be defined as $y = x^3 + 1$ over $\mathbb{F}_q$, where q is 512 bits. In addition, the programming languages (e.g., C, C++, Java, Python) and cryptographic libraries (e.g., Miracl, PBC, NTL) can be used to implement P^{2}SM. We follow the TLS handshake protocol to describe the information exchange between two entities as shown in Fig. 5. Based on the TLS protocol, we define the format of all messages exchanged among smart meters, collectors, the operation center and customers, and each algorithm in P^{2}SM is deployed on the corresponding entity. Specifically, the report generation algorithm is deployed on smart meters, the report aggregation algorithm should be implemented on collectors, the report reading algorithm is executed on the operation center, and the dynamic billing algorithm is executed among collectors, the operation center and customers to generate the daily bills. In summary, we can utilize the algorithms in P^{2}SM to replace the encryption, signature, and authentication algorithms in TLS to achieve

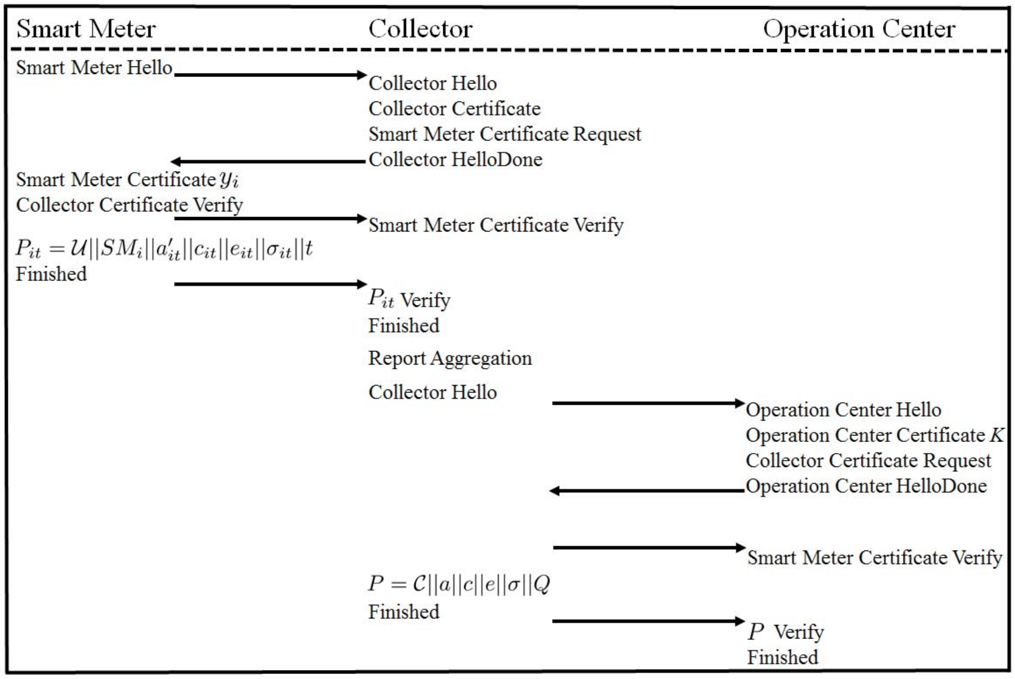

Smart Meter
Collector
Operation Center
Smart Meter Hello
Collector Hello
Collector Certificate
Smart Meter Certificate Request
Collector HelloDone
Smart Meter Certificate y_i
Collector Certificate Verify
Smart Meter Certificate Verify
$P_{it} = \mathcal{U}||SM_i||a'_{it}||c_{it}||e_{it}||\sigma_{it}||t$
Finished
P_{it} Verify
Finished
Report Aggregation
Collector Hello
Operation Center Hello
Operation Center Certificate K
Collector Certificate Request
Operation Center HelloDone
Smart Meter Certificate Verify
$P = \mathcal{C}||a||c||e||\sigma||Q$
Finished
P Verify
Finished

Fig. 5. Implementation of P^{2}SM.

end-to-end security, data aggregation and dynamic billing for smart metering.

VII. Conclusion

In this paper, we have introduced a new security model to formally define the misbehavior of hacked collectors and have proposed a privacy-preserving smart metering scheme to achieve end-to-end security and high communication efficiency in smart grid. P^{2}SM not only allows collectors to aggregate authentication messages, meter readings and signatures to reduce communication overhead and preserve the privacy of customers, but also prevents a misbehaving collector from corrupting power consumption reports. Further, the collector is able to generate verifiable daily electricity bills from individual consumption reports based on dynamic prices for customers. We have proved the security of P^{2}SM and evaluated its performance through the comparison with the existing schemes. P^{2}SM is a secure and efficient communication protocol that can replace the TLS protocol to achieve secure smart metering in smart grid. For the future work, we will design a strong privacy-preserving smart metering scheme from local differential privacy against curious operation centers in smart grid.

References

[1] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid—The new and improved power grid: A survey," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 4, pp. 944–980, 4th Quart., 2012.

[2] X. Li *et al.*, "Securing smart grid: Cyber attacks, countermeasures, and challenges," *IEEE Commun. Mag.*, vol. 58, no. 8, pp. 38–45, Aug. 2012.

[3] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin, "Private memoirs of a smart meter," in *Proc. BuildSys*, Zürich, Switzerland, 2010, pp. 61–66.

[4] *The Smart Grid and Privacy*, Electron. Privacy Inf. Center, Washington, DC, USA, 2015.

[5] *IEC TC57 WG15: Security Standards for the Power System Information Infrastructure*, IEC Standard 62351, Jun. 2012.

[6] T. Dierks and E. Rescorla, "The transport layer security (TLS) protocol-version 1.2," NFC Working Group, Aug. 2008.

[7] *Government of Ontario IT Standard: Advanced Metering Infrastructure*, document 51, Govern. Ontario, Toronto, ON, Canada, 2007.

[8] 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 Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1621–1631, Sep. 2012.

[9] H.-Y. Lin, W.-G. Tzeng, S.-T. Shen, and B.-S. P. Lin, "A practical smart metering system supporting privacy preserving billing and load monitoring," in *Proc. ACNS*, Singapore, 2012, pp. 544–560.

[10] C. Rottondi, G. Verticale, and A. Capone, "Privacy-preserving smart metering with multiple data consumers," *Comput. Netw.*, vol. 57, no. 7, pp. 1699–1713, 2013.

[11] T. Dimitriou and G. Karame, "Privacy-friendly tasking and trading of energy in smart grids," in *Proc. ACM SAC*, Coimbra, Portugal, 2013, pp. 652–659.

[12] C. Rottondi, M. Savi, G. Verticale, and C. Krauß, "Mitigation of peer-to-peer overlay attacks in the automatic metering infrastructure of smart grids," *Security Commun. Netw.*, vol. 8, no. 3, pp. 343–359, 2014.

[13] H. Krawczyk and T. Rabin, "Chameleon signatures," in *Proc. NDSS*, 2000, pp. 143–154.

[14] G. Ateniese, K. Fu, M. Green, and S. Hohenberger, "Improved proxy re-encryption schemes with applications to secure distributed storage," in *Proc. NDSS*, 2005, pp. 29–43.

[15] H. Shacham and B. Waters, "Compact proofs of retrievability," in *Proc. Asiacrypt*, 2008, pp. 90–107.

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

[17] K. Ohara, Y. Sakai, F. Yoshida, M. Iwamoto, and K. Ohta, "Privacy-preserving smart metering with verifiability for both billing and energy management," in *Proc. AsiaPKC*, Kyoto, Japan, 2014, pp. 23–32.

[18] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *Proc. EUROCRYPT*, Prague, Czech Republic, 1999, pp. 223–238.

[19] D. Boneh, E.-J. Goh, and K. Nissim, "Evaluating 2-DNF formulas on ciphertexts," in *Proc. TCC*, Cambridge, MA, USA, 2005, pp. 325–341.

[20] D. Boneh, B. Lynn, and H. Shacham, "Short signatures from the Weil pairing," in *Proc. Asiacrypt*, 2001, pp. 514–532.

[21] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," *IEEE Trans. Inf. Theory*, vol. 31, no. 4, pp. 469–472, Jul. 1985.

[22] K. Kursawe, G. Danezis, and M. Kohlweiss, "Privacy-friendly aggregation for the smart-grid," in *Proc. PETS*, Waterloo, ON, Canada, 2011, 175–191.

[23] W. Jia, H. Zhu, Z. Cao, X. Dong, and C. Xiao, "Human-factor-aware privacy-preserving aggregation in smart grid," *IEEE Syst. J.*, vol. 8, no. 2, pp. 598–607, Jun. 2014.

[24] J. Won, C. Y. T. Ma, D. K. Y. Yau, and N. S. V. Rao, "Privacy-assured aggregation protocol for smart metering: A proactive fault-tolerant approach," *IEEE/ACM Trans. Netw.*, vol. 24, no. 3, pp. 1661–1674, Jun. 2016.

[25] J. Ni *et al.*, "Differentially private smart metering with fault tolerance and range-based filtering," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2483–2493, Sep. 2017.

[26] Z. Zhang, Z. Qin, L. Zhu, J. Weng, and K. Ren, "Cost-friendly differential privacy for smart meters: Exploiting the dual roles of the noise," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 619–626, Mar. 2017.

[27] J. Ni, K. Alharbi, X. Lin, and X. Shen, "Security-enhanced data aggregation against malicious gateways in smart grid," in *Proc. Globecom*, San Diego, CA, USA, 2015, pp. 1–6.

[28] H. J. Jo, I. S. Kim, and D. H. Lee, "Efficient and privacy-preserving metering protocols for smart grid systems," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1732–1742, May 2016.

[29] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. ACM CCS*, Chicago, IL, USA, 2009, pp. 21–32.

[30] J. Hu and A. V. Vasilakos, "Energy big data analytics and security: Challenges and opportunities," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2423–2436, Sep. 2016.

[31] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1630–1638, Jul. 2017.

[32] J. Chen and A. Abur, "Placement of PMUs to enable bad data detection in state estimation," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1608–1615, Nov. 2006.

[33] Q. Yang *et al.*, "On false data-injection attacks against power system state estimation: Modeling and countermeasures," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 3, pp. 717–729, Mar. 2014.

[34] A. Ashok, M. Govindarasu, and V. Ajjarapu, "Online detection of stealthy false data injection attacks in power system state estimation," *IEEE Trans. Smart Grid*, to be published.

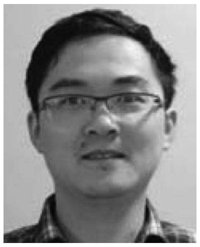
[35] Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2505–2516, Sep. 2017.

[36] D. Derler and D. Slamanig, "Key-homomorphic signatures and applications to multiparty signatures and non-interactive zero-knowledge," Cryptol. ePrint Arch., Tech. Rep. 792, 2016. [Online]. Available: https://eprint.iacr.org/2016/792.pdf

[37] J. M. Pollard, "Kangaroos, monopoly and discrete logarithms," *J. Cryptol.*, vol. 13, no. 4, pp. 437–447, 2000.

[38] B. Libert and J.-J. Quisquater, "Improved signcryption from $q-$Diffie–Hellman problem," in *Proc. SCN*, Amalfi, Italy, 2004, pp. 220–234.

[39] C.-H. Li and J. Pieprzyk, "Conference key agreement from secret sharing," in *Proc. ACISP*, Wollongong, NSW, Australia, 1999, pp. 64–76.



Jianbing Ni (S'16) received the B.E. and M.S. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 2011 and 2014, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His research interests are applied cryptography and information security, with current focus on cloud computing, smart grid, and Internet of Things.

Kuan Zhang (S'13–M'17) received the B.Sc. degree in communication engineering and the M.Sc. degree in computer applied technology from Northeastern University, China, in 2009 and 2011, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Canada, in 2016, where he was also a Post-Doctoral Fellow with the Broadband Communications Research Group, Department of Electrical and Computer Engineering, University of Waterloo. Since 2017, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, University of Nebraska-Lincoln, USA. His research interests include security and privacy for mobile social networks, e-healthcare systems, cloud computing, and cyber physical systems.

Xiaodong Lin (M'09–SM'12–F'17) received the Ph.D. degree in information engineering from the Beijing University of Posts and Telecommunications, Beijing, China, and the Ph.D. degree in electrical and computer engineering (with Outstanding Achievement in Graduate Studies Award) from the University of Waterloo, Waterloo, ON, Canada. He was an Associate Professor of information security with the Faculty of Business and Information Technology, University of Ontario Institute of Technology, Canada. He is currently an Associate Professor with the Department of Physics and Computer Science, Wilfrid Laurier University, Waterloo. His research interests include wireless network security, applied cryptography, computer forensics, software security, and wireless networking and mobile computing.

Xuemin (Sherman) Shen (M'97–SM'02–F'09) received the B.Sc. degree from Dalian Maritime University, China, in 1982 and the M.Sc. and Ph.D. degrees from Rutgers University, NJ, USA, in 1987 and 1990, respectively, all in electrical engineering. He is a University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research focuses on resource management in interconnected wireless/wired networks, wireless network security, social networks, smart grid, and vehicular ad hoc and sensor networks. He was a recipient of the Excellent Graduate Supervision Award in 2006, and the Outstanding Performance Award in 2004, 2007, 2010, and 2014 from the University of Waterloo, the Premier's Research Excellence Award in 2003 from the Province of Ontario, Canada, the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo, and the Joseph LoCicero Award from the IEEE Communications Society. He served as the Technical Program Committee Chair/Co-Chair for IEEE Globecom'16, Infocom'14, IEEE VTC'10 Fall, and Globecom'07, the Symposia Chair for IEEE ICC'10, the Tutorial Chair for IEEE VTC'11 Spring and IEEE ICC'08, the General Co-Chair for ACM Mobihoc'15, Chinacom'07, and QShine'06, and the Chair for IEEE Communications Society Technical Committee on Wireless Communications, and P2P Communications and Networking. He also serves/served as the Editor-in-Chief for the *IEEE Internet of Things Journal*, IEEE Network, *Peer-to-Peer Networking and Application*, and *IET Communications*; the Founding Area Editor of the IEEE Transactions on Wireless Communications; an Associate Editor of the IEEE Transactions on Vehicular Technology, *Computer Networks*, and *ACM/Wireless Networks*. He is a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, and a Distinguished Lecturer of IEEE Vehicular Technology Society and Communications Society.