Hindawi
Security and Communication Networks
Volume 2019, Article ID 9637578, 14 pages
https://doi.org/10.1155/2019/9637578

Research Article

EPLC: An Efficient Privacy-Preserving Line-Loss Calculation Scheme for Residential Areas of Smart Grid

Yinqiao Xiong,[1,2] **Peidong Zhu**,[2] **Zhizhu Liu**,[1] **Hui Yin**,[2] **and Tiantian Deng**[1,2]

[1] *College of Computer, National University of Defense Technology, Changsha 410073, China*
[2] *College of Electronic Information and Electrical Engineering, Changsha University, Changsha 410022, China*

Correspondence should be addressed to Peidong Zhu; pdzhu@nudt.edu.cn

Received 26 October 2018; Accepted 7 March 2019; Published 27 March 2019

Guest Editor: Esther Palomar

Recently, smart grid is considered as the next generation of power grid by introducing information and communication technologies. Line-loss is an important synthetic indicator which can directly reflect the energy efficiency and power management level of smart grid enterprises. In order to obtain all residential areas, line-loss requires obtaining electricity consumption of each user. However, data about users' electricity consumption could reveal sensitive information; a sophisticated adversary can use some data analysis methods to deduce economic situation, habits, lifestyles, etc. In order to solve the problem, we propose an Efficient Privacy-preserving scheme for Line-loss Calculation, named EPLC. In our scheme, a data item is reading from one smart meter which implies the energy consumption in a time period of the user who owns it, and each user lives in a residential area. For each user, we encrypt user's data based on Paillier cryptosystem by using two Horner parameters, by leveraging homomorphism, and each residential area gateway calculates relevant data about corresponding line-loss and control center hides the area-level polynomial into the final output for representing line-loss of all residential areas which are both in the form of ciphertext. Finally, we can still recover each residential area line-loss with possessing private keys and Horner parameters. Moreover, EPLC adopts the batch verification technique to lower authentication cost. Finally, our analysis indicates that EPLC is not only efficient but also can protect individual user's electricity consumption privacy, and the flexibility and expansibility of EPLC are very suitable for smart grid.

1. Introduction

In the modern world, electrical energy plays a crucially important role in economic, social, and industrial development of all nations and regions. Smart grid, defined by the US Department of Energy [1], is considered to be the next generation of power grid infrastructure by integrating information and communication technologies (ICT) and can real-time monitor and control the physical processes of power system to constantly heighten the electrical energy using efficiency and optimize services to adjust electrical energy supply to meet demand, so smart grid is characterised by high intelligence, efficiency, reliability, economic behavior, and security. Figure 1 shows the architecture of smart grid provided by National Institute of Standards and Technology (NIST) [2], which contains seven domains: generation domain, generating electrical energy; transmission domain, transmitting electrical energy to and from the distribution; distribution domain, distributing electrical energy to and from customers; customers domain, users or home area networks (HANs); operation domain, managing the electricity flow; market domain, managing the electricity market in the smart grid; service provider domain, managing all the third-party operations.

Smart meter is a kind of intelligent instrument with significant capabilities for two-way communication, measuring and reporting electricity consumption in near real-time (15 minutes) periodically [3, 4], and also as one kind of the key components for realization of smart grid deployed at users, voltage transformers, and wherever needed. Accordingly, smart meter is set to become one of the most important sources of near real-time electrical energy flows data in smart grid. Based on the detailed view on electrical energy flows, through near real-time state estimation, not only the

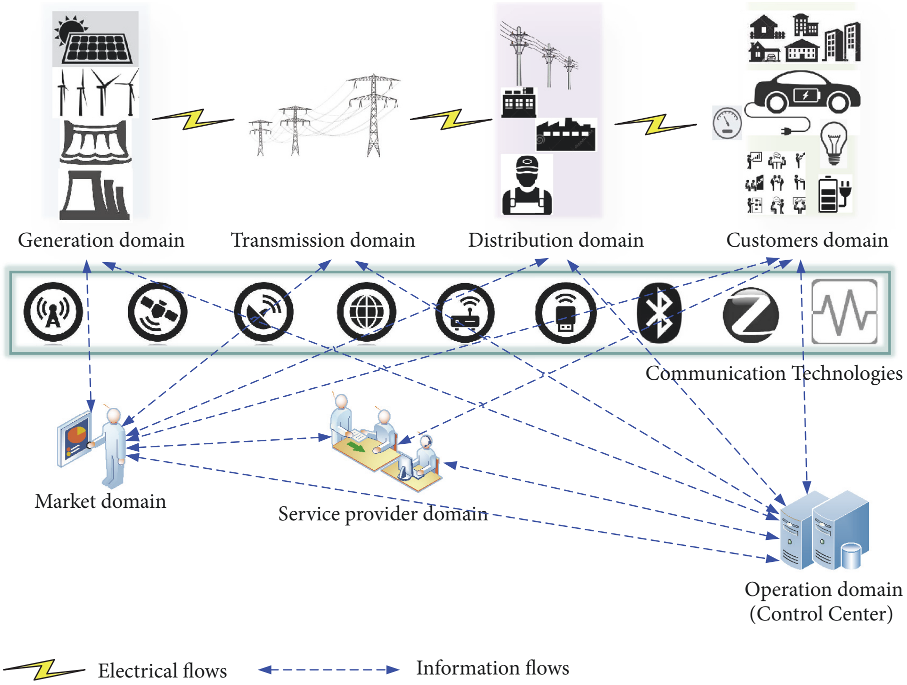


Figure 1: Smart grid architecture.

monitoring of smart grid's performance, but also optimization of supply, distribution, and consumption can be executed synchronously well.

Line-loss, the decrease of electrical energy from the source to destination due to inherent inefficiencies and defects in smart grid, is an important synthetic indicator to reflect the management level and is also a main content of assessment for performance level of smart grid enterprises [5]. A variety of factors such as resistance effect (variable losses, with the increasing of current), electromagnetism influence (fixed losses, with the increasing of voltage), management factors (electric larceny, metering error, and electricity leakage), etc. are the causes of the line-loss. Especially for residential areas which are located in the grid terminal and the most bottom of administrative level, they generally consist of various kinds of equipment and directly associated with all kinds of users, so line-loss such as electric larceny, leakage, and etc. is much more likely to happen. As a result, we only consider the line-loss of residential areas in our scheme.

Fortunately, based on the idea of electrical energy quantity [6] for line-loss calculation, the basic method is fairly simple by considering the line-loss as distributed sale electric quantity noncorrespondence [7, 8], which can extrapolate line-loss for most environment or variables, so the process of line-loss calculation for each residential area only needs to do some sums and subtractions over the corresponding data of near real-time electricity consumption recorded by smart meters. But unfortunately these data can directly reflect the privacy of users [9–13] such as the occupancy of a household, and an sophisticated adversary $\mathscr{A}$ can leverage some data mining algorithms by some clever tricks to infer the lifestyle habits, economic status, etc. [3, 4, 14], and even users' interests as well [15]. Worse yet, some opportunities for criminal purpose will be provided too [16].

In general, smart meter has its own secure components [17], and symmetric cryptosystem is introduced [11, 18], after completion symmetric encryption operation. Although all the sensitive and private data measured by smart meter achieve cryptographic storages as well as communications, many utilities such as the line-loss calculation mentioned above and data that are encrypted need to be decrypted before they can be used. Therefore, there are still considerable risks for exposure of users' sensitive information by inside attack [19].

EPLC is the scheme mainly focusing on preserving privacy of users while calculating the line-loss by the method mentioned above of each residential area. To preserve privacy of users from internal and external attacks, the communication channels, gateways, and servers are not fully trusted. All sensitive data should be encrypted against eavesdropping before sending; all receivers should verify the ciphertext received to catch tampering before processing. The line-loss calculation and the storage both depend on ciphertext to resist the attack from the inside such as administrator of servers or gateways, and all the detailed elaboration will be made in a later section.

The remainder of our paper is organized as follows: in Section 2, overview of related works is provided; Section 3 elaborates on system model, security requirements, and design goal; we introduce the bilinear pairing, Paillier cryptosystem, and Horner's rule as the preliminaries of EPLC in Section 4; Section 5 presents the process of EPLC scheme

Table 1: Main notations.

Notation	Description
CC, VT, SM	Control Center, Voltage Transformer, Smart Meter
DA, RA	District Area, Residential Area
D_i	The ith district area of system model
R_{ij}	The jth residential area of the ith district area
U_{ijk}	The kth user in the R_{ij}
$U_{ij\tau}$	The VT which is equipped in the R_{ij}
RAGW$_{ij}$	The RAGW which belongs to the R_{ij}

step by step; security analysis and performance evaluation are described in Sections 6 and 7, respectively; finally, a conclusion is given in Section 8.

2. Related Work

Recently, the leakage of personalized privacy in smart grid from data of smart meters which demonstrate electricity consumption of users becomes a potential problem. For preserving users' privacy, predecessors have proposed several approaches.

Y. Sun et al. [20] hide household load in the data of smart meters by leveraging existed thermal appliances and energy storage units to protect privacy of users. The study [21] obfuscates the real electricity consumption based on masking. Based on the differential privacy-preserving technique [22] and Boneh-Goh-Nissim cryptosystem [23], the scheme [24] uses Laplace noise in the form of ciphertext to protect the privacy of users in the honest-but-curious model.

For privacy-preserving in aggregation, super-increasing sequences and Horner's ruler are introduced to structure multidimensional data of smart meters, respectively [10, 25]. The scheme [26], named privacy-preserving multisubset aggregation (PPMA), divides users electricity consumption data into different subsets which represented different ranges before aggregation for improving the efficiency. Reference [27] can aggregate hybrid IoT devices data into one with resisting against false data injection attack; the scheme [28] performs aggregation with privacy-preserving and fault tolerance and recovers the private key by Shamirs secret-sharing scheme [29]. All works [10, 25–28] implement aggregation of smart meters' data as well as keeping the privacy of users are based on the homomorphic property of Paillier cryptosystem [30]. Additionally, privacy-preserving aggregation is performed by sharing the session keys of users with wireless [31]; fully homomorphic encryption (FHE) and secure multiparty computation are leveraged for achieving users privacy preserving in secure in-network data aggregation by smart grid V2G networks [32], W. Han et al. proposed an integrated privacy-preserving data management architecture (IP^{2}DM) [33] which achieves anonymous data aggregation by partial homomorphic encryption [34] for privacy-preserving data management.

According to the line-loss calculation method mentioned above, this paper focuses on efficient line-loss calculation of residential areas based on data of smart meters as well as avoiding any leaks of private information of users.

3. System Model, Security Requirements, and Design Goal

In this section, we formalize the system model and security requirements, identify our design goals, and refer to Table 1 for listing some notations to represent the main entities and definitions in system model.

3.1. System Model. In our system model, our focus is on how to get the line-loss in residential areas correctly and efficiently while keeping users' privacy. As shown in Figure 2, on the one hand by the aspect of electricity, the electrical energies generated from power plants, solar panels, windmills, and etc. are transmitted through transmission and distribution to each voltage transformer (VT) of residential areas and then transmit to users by using electrical lines of the residential area (RA) to meet needs, and some RAs form a district area (DA) generally; on the other hand, according to the information technology, there are four types of entities: Control Center (CC), Residential Area Gateway (RAGW), Home Area Network (HAN or user), and Smart Meter (SM). In general, there is only one CC in the system, according to the needs of self-business management, the CC manages at most m DAs $\mathbb{D} = \{D_1, D_2, \ldots, D_k, \ldots, D_m\}$, and each DA contains at most n RAs; the RAs in the ith DA D_i could be described as $\mathbb{R} = \{R_{i1}, R_{i2}, \ldots, R_{ik}, \ldots, R_{in}\}$. Also, each RA has at most s HANs or residential users, taking a typical RA R_{ij} as an example, as shown in Figure 3, which means the jth $(1 \leq j \leq n)$ RA of the ith $(1 \leq i \leq m)$ DA managed by the CC in the system, the R_{ij} comprises at most s residential users $\mathbb{U} = \{U_{ij1}, U_{ij2}, \ldots, U_{ijk}, \ldots, U_{ijm}\}$, and the notation $U_{ijk} \in \mathbb{U}$ means the kth user of the jth RA in the ith DA. $(1 \leq i \leq m,\ 1 \leq j \leq n \ and\ 1 \leq k \leq s)$.

The nearly real-time (e.g., 15 minutes) electricity consumption of users and RAs can be recorded automatically and respectively by corresponding smart meters, which are equipped in each HAN (User) and VT of RA. RAGW is a powerful computing resource and primarily completes three functions: authentication, line-loss calculation for corresponding RA, and relaying. Firstly, RAGW can perform some authentication operations to guarantee the received data's authenticity and integrity; secondly, by getting the data

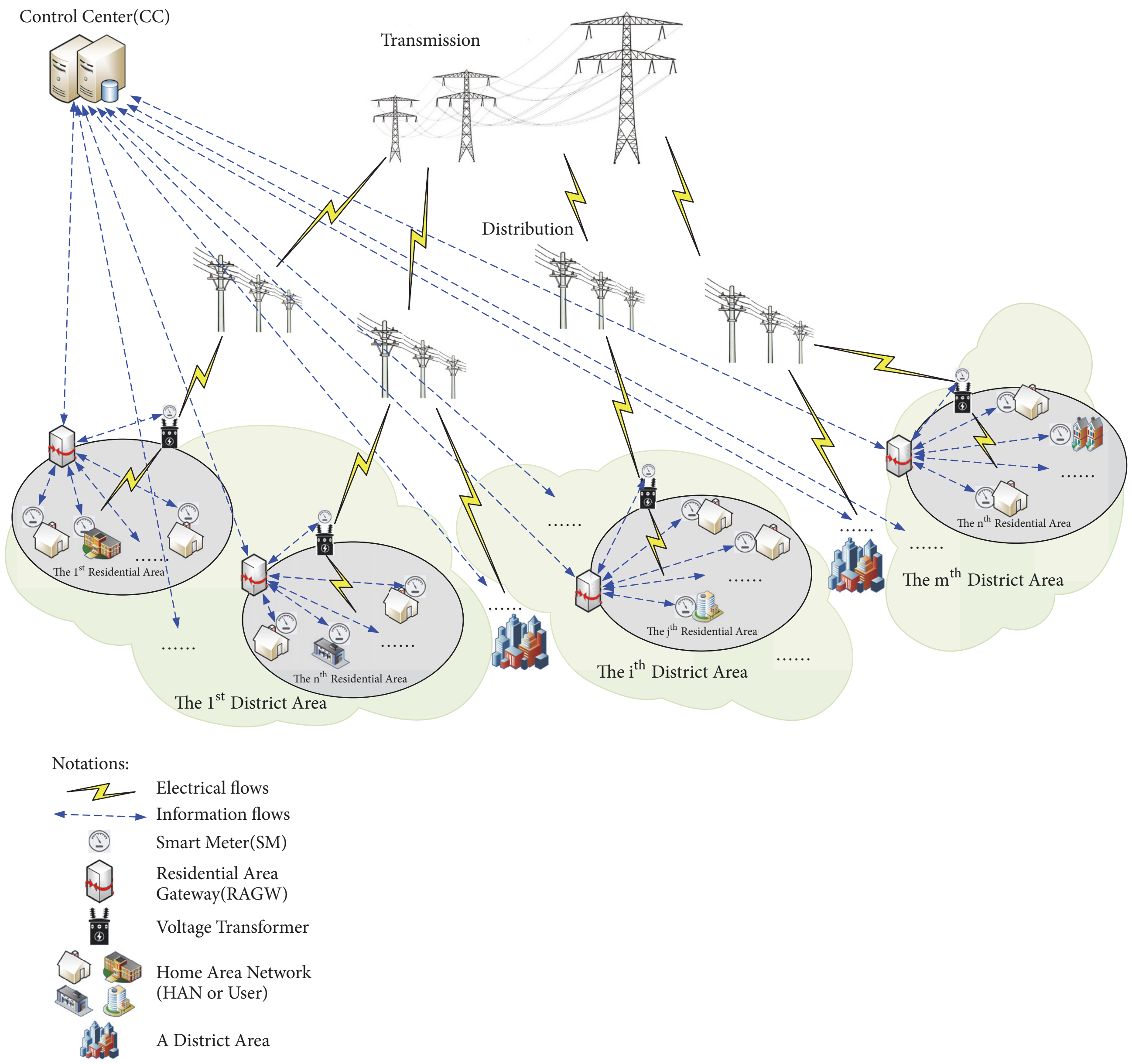


Figure 2: System model.

(ciphertext) from smart meters of users and VT of corresponding RA, RAGW can calculate the line-loss ciphertext of the RA; thirdly, the relaying component is responsible for forwarding the line-loss ciphertext of corresponding RA to CC. According to business need, CC is responsible for receiving the reports (ciphertext) from all RAGWs and decrypting each line-loss of RAs, which can help itself get the real-time situation awareness and produce some responses.

As mentioned before, according to Figure 3, all electricity consumption of users in one RA comes from the VT of the RA. In other words, all the electrical energies consumed by users of the RA can be recorded periodically (e.g., 15 minutes) by the smart meter equipped on the VT. However, the data of the VT's smart meter record not only the electricity consumption of all users in the RA, but also the line-loss of the RA, such as loss of the electrical lines and other types of equipment, equipment failures, and even electric larceny in the process of energies transmission as shown in red dash dotted box of Figure 3. To get the line-loss of each residential area, firstly we make the aggregation by collecting all users' electricity consumption of the corresponding residential area, respectively; secondly, we calculate the difference by subtracting the aggregation result from the corresponding reading of VT's smart meter.

Typically, considering the communication model in one residential area RA as shown in Figure 3, the number of users is limited, the distances between the two sides of the communications are short and there are not too much electromagnetic interferences, so the communications between users $U_{ijk} \in \mathbb{U}$ and RAGW of the RA can use relatively

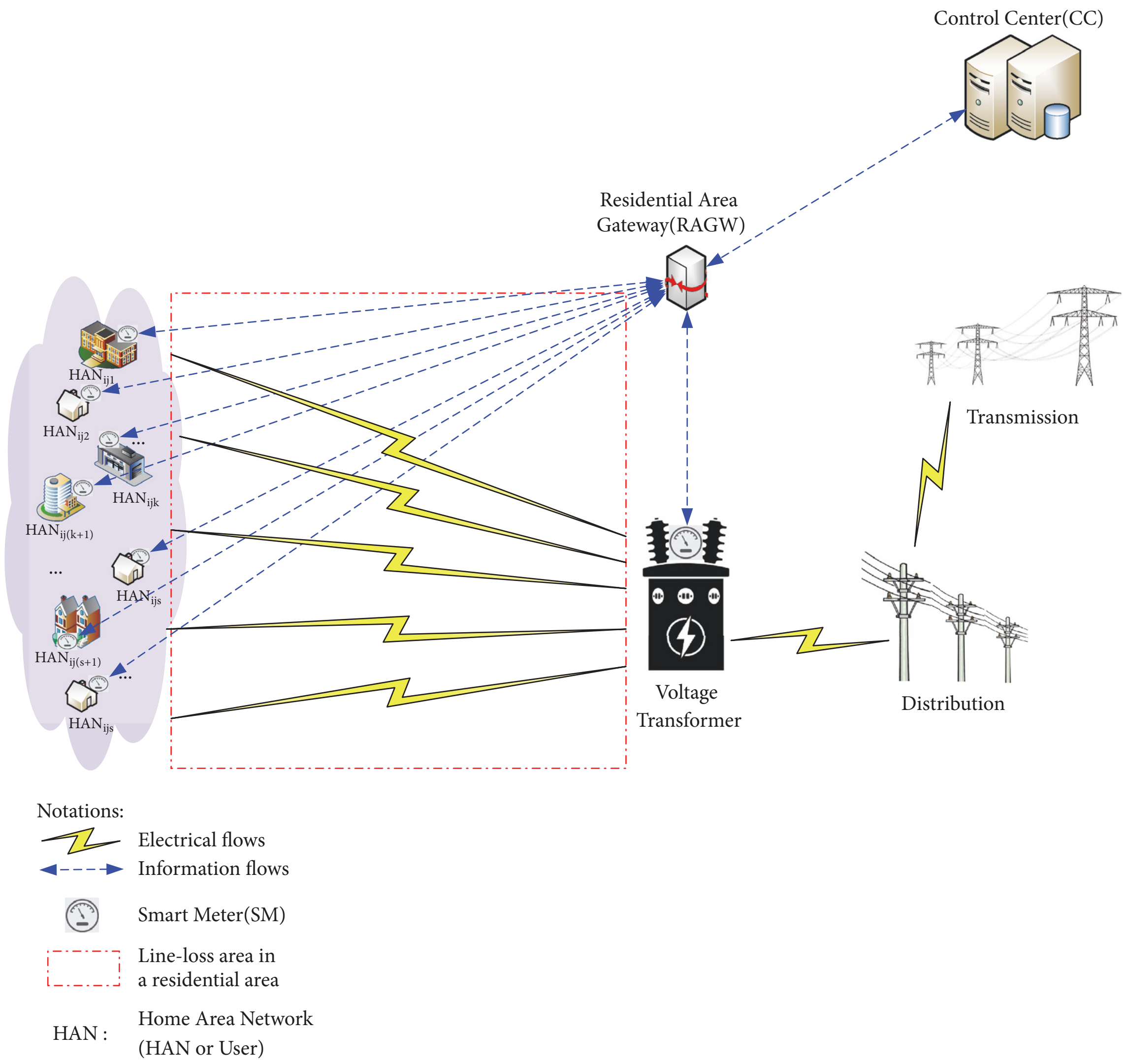


Figure 3: System model of a residential area.

inexpensive wireless technology, e.g., WiFi technology. On the other hand, according to Figure 2, since the distances between the RAGWs and CC are far away, the communications between them are through the use of some wired links with high bandwidth and low delay, like optical fiber.

3.2. Security Requirements. Security is crucial for our scheme. In our security model, we consider the CC fully trusted, and RAGWs and users follow the honest-but-curious model. Although RAGW will execute correctly according to design, keeping all data from smart meters and all intermediate computational results will lead to a huge threat of users' privacy leakage. Users will not drop or distort any source data spitefully and keep the system running correctly. However, another potential risk for privacy leakage is using collusion attack to infer other users' electricity consumption. There might exist an external adversary $\mathscr{A}$ in the system, who can eavesdrop not only residential users' but also RAGWs' reports on different kinds of communication channels, like wireless technology in RA and wired links between CC and RAGW. More seriously, a powerful adversary $\mathscr{A}$ even could intrude in the databases of RAGWs and CC to steal some private and personal-sensitive data of electricity consumption or put forward to launch some active attacks to compromise the data integrity. As is clear from the above descriptions, to avoid being sniffed and to detect malicious actions which are both by the adversary $\mathscr{A}$, we should achieve the following security requirements in our scheme:

(i) *Confidentiality*. In our system, for protecting privacy of individual users of residential areas against malicious actions of adversary $\mathscr{A}$, even if eavesdropping occurs on communication channels and data in databases of CC or RAGWs have been stolen, adversary $\mathscr{A}$ also can neither obtain any information of individual users' electricity energy consumption, nor identify or infer any other users' privacy information.

(ii) *Authentication and Data Integrity*. Notice that RAs are in public places, a skilled adversary $\mathscr{A}$ can easily hack into the communication system, then forge or modify reports which are sent by legal residential users. So we must authenticate encrypted reports which are really sent by their corresponding legal residential users and have not been tampered during the transmission.

3.3. Design Goal. Under our system model and for achieving the aforementioned security requirements, our design goal is to develop an efficient, privacy-preserving line-loss calculation scheme. Particularly, the scheme achieved the following three objectives:

(i) *Security and Reliability*. The security and reliability are the most essential goals in our scheme. Without the security and reliability, the privacy information about users' real-time electricity consumption will be leaked or even tampered, after which error results will be generated in processing the tampered information and then sent to CC. If the error results are applied in the planning, operation or analysis fields of power system, it will cause collapse for the worst. So our proposed scheme should guarantee the authentication, confidentiality, and integrity simultaneously.

(ii) *High Efficiency*. Consider the real-timeness requirement and characteristics of communication architecture in smart grid system. For example, the communication channels in RA always use wireless technology, which is featured with low-bandwidth and high-delay compared to wired technology, and smart meters have limited computing ability, memory, and so on. So our proposed scheme should reduce communication cost and improve the efficiency of the line-loss calculation processing.

(iii) *Good Flexibility*. Even though the numbers of users, RAs, or DAs varied, the business logic changed constantly, etc., the proposed scheme still can carry out flexible line-loss calculation for each RA very well.

4. Preliminaries

Based on the bilinear pairing technique [35], the Paillier cryptosystem [30], and Horner's rule [36], we propose the EPLC scheme.

4.1. The Bilinear Pairing. Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be two additive cyclic groups of the same large prime number q, P be a generator of group $\mathbb{G}_1$, and let $e : \mathbb{G}_1 \times \mathbb{G}_1 \longrightarrow \mathbb{G}_2$ be a bilinear map. The bilinear pairing contains the following properties.

(i) *Bilinearity*. $e(aP, bQ) = e(P, Q)^{ab} \in \mathbb{G}_2$ for $\forall P,\ Q \in \mathbb{G}_1$ and $\forall a,\ b \in \mathbb{Z}_q^*$.

(ii) *Nondegeneracy*. $e(P, P) \neq 1_{\mathbb{G}_2}$.

(iii) *Computability*. For any $P,\ Q \in \mathbb{G}_1$, there exists an efficient polynomial time algorithm to compute $e(P, Q)$ and the group operation in $\mathbb{G}_1$ is also efficiently computable.

(iv) *Symmetry*. The map e is symmetric: $e(aP, bQ) = e(P, Q)^{ab} = e(bP, aQ)$.

Bilinear Pairing Generation Algorithm. A bilinear parameter generator $\mathscr{G}en$ is a probabilistic algorithm, which takes a security parameter κ as input and then outputs a five-tuple $(q, P, \mathbb{G}_1, \mathbb{G}_2, e)$; in the tuple, the q is a κ-bit prime number, the $P \in \mathbb{G}_1$ is a generator, and the $e : \mathbb{G}_1 \times \mathbb{G}_1 \longrightarrow \mathbb{G}_2$ is a bilinear map which has the above properties.

Computational Diffie-Hellman (CDH) Problem [37]. By given two elements S_1 and S_2 from group $\mathbb{G}_1$, define $\mathsf{DH_P}(S_1, S_2) \overset{def}{=} abP$, when $S_1 = aP$ and $S_2 = bP$, then

$$\mathsf{DH_P}\left(S_1, S_2\right) = abP = aS_2 = bS_1 \tag{1}$$

The definition of *CDH* problem is given randomly chosen S_1 and S_2; then compute $\mathsf{DH_P}(S_1, S_2)$.

4.2. The Paillier Cryptosystem. Just as the Goldwasser-Micali, RSA, and Rabin encryption schemes, the Paillier cryptosystem is also based on the hardness of factoring a composite number N which is a product of two prime numbers, while the more specially and importantly is that the Paillier cryptosystem possesses unique *homomorphic* characteristic property, and its efficiency is almost the same with RSA and Rabin but higher than Goldwasser-Micali.

The Paillier cryptosystem utilizes the group $\mathbb{Z}_{N^2}^*$, which is isomorphic to $\mathbb{Z}_N \times \mathbb{Z}_N^*$ with isomorphism $f : \mathbb{Z}_N \times \mathbb{Z}_N^* \longrightarrow \mathbb{Z}_{N^2}^*$

$$f(a, b) = \left[(1 + N)^a \cdot b^N \bmod N^2\right] \tag{2}$$

where $N = pq$, p and q are both big prime numbers of the same length with different values, $a \in \mathbb{Z}_N$, $b \in \mathbb{Z}_N^*$, $f(a, b) \in \mathbb{Z}_{N^2}^*$, and $gcd(N, \phi(N)) = 1$, the order of $(1 + N)$ in $\mathbb{Z}_{N^2}^*$ is N because for an integer a with $0 \leq a \leq N$, and the equation $(1 + N)^a = (1 + aN) \bmod N^2$ will always be true.

Algorithms of Paillier Cryptosystem

(i) *Key Generation.* This cryptosystem takes a security parameter ξ as input and then chooses two big prime numbers p and q of the same length ξ, output $N = pq$, and $\lambda = lcm(p - 1, q - 1)$; define a function $L(u) = (u - 1)/N$, then choose $g = (1 + N) \in \mathbb{Z}_{N^2}^*$, and calculate $\mu = (L(g^\lambda \bmod N^2))^{-1} \bmod N$. Finally distribute the public keys $pk = (N, g)$ and the corresponding private keys $sk = (\lambda, \mu)$.

(ii) *Encryption.* For each user's message $m \in \mathbb{Z}_N$, a random number $r \in \mathbb{Z}_N^*$ was chosen by the user; then generate the ciphertext $c = E(m) = g^m \cdot r^N \bmod N^2$.

(iii) *Decryption.* When the user get a ciphertext $c \in \mathbb{Z}_{N^2}^*$, deciphering it by $D(c) = L(c^{\lambda \bmod N^2}) \cdot \mu \bmod N$, the corresponding message m can be got.

4.3. Horner's Rule. The Horner algorithm is a fast algorithm for computing polynomials, which was named for William George Horner who is an English mathematician. According

Table 2: Main parameters.

Parameter	Description
$\kappa, q, P, \mathbb{G}_1, \mathbb{G}_2, e$	Parameters of Bilinear Pairing
$\xi, N, p, q, g, \lambda, \mu$	Parameters of Pallier Cryptosystem
x_{ij}, Y_{ij}	The private key and public key of RAGW_{ij}
x_{ijk}, Y_{ijk}	The private key and public key of U_{ijk}
$x_{ij\tau}, Y_{ij\tau}$	The private key and public key of $\mathrm{VT}(U_{ij\tau})$
Υ_1, Υ_2	Two common factors in line-loss calculation
m, n, s	The maximum numbers of DAs, RAs of each DA and users in each RA respectively
Θ	The maximum number of line-loss in each RA
$\mathscr{N}^{DA}$	After initialization, the number of DAs
$\mathscr{N}_i^{RA}$	After initialization, the number of RAs in the D_i
$\mathscr{N}_{ij}^{U}$	After initialization, the number of users in the R_{ij}
$LLD\_C_{ij\mathscr{T}}$	The ciphertext of line-loss in R_{ij} at time $\mathscr{T}$
$LLD\_P_{ij\mathscr{T}}$	The plaintext of line-loss in R_{ij} at time $\mathscr{T}$
$LLD\_C_{\mathscr{T}}$	All the line-loss of the system at time $\mathscr{T}$ in the form of cipher as a number
$LLD\_P_{\mathscr{T}}$	All the line-loss of the system at time $\mathscr{T}$ in the form of polynomials as a number

to a parameter Υ, Horner's rule will actually be able to turn any polynomial expressed as $p(\Upsilon) = a_0 + a_1\Upsilon + \cdots + a_k\Upsilon^k + \cdots + a_n\Upsilon^n$ into another form $p(\Upsilon) = a_0 + \Upsilon(a_1 + \Upsilon(a_2 + \cdots + \Upsilon(a_{n-1} + a_n\Upsilon)))$. After transformation, calculating the polynomial only needs n multiplications and n additions; obviously it is more efficient than before.

Particularly, given a set of data $\mathbb{V} = \{v_0, v_1, \ldots, v_k, \ldots, v_n\}$ and $\Upsilon > \max\{v_0, v_1, \cdots, v_k, \ldots, v_n\}$, then we can construct a polynomial $p(\Upsilon) = v_0 + v_1\Upsilon + \cdots + v_k\Upsilon^k + \cdots + v_n\Upsilon^n = v_0 + \Upsilon(v_1 + \Upsilon(v_2 + \cdots + \Upsilon(v_{n-1} + v_n\Upsilon)))$, then all the information of the set $\mathbb{V}$ is interpreted by the $p(\Upsilon)$ as a number. After that, if only we know the value of the $p(\Upsilon)$ and Υ, retrieving the $\mathbb{V}$ by n exact divisions and modulo operations, respectively, is going to be easy.

5. Our Proposed EPLC Scheme

In this section, an efficient privacy-preserving line-loss calculation scheme for residential areas of smart grid is proposed, which is made up of the following four parts: system initialization, user report generation, privacy-preserving line-loss calculation, and decryption of line-loss. Following the description of system model, we assume the numbers of DAs, RAs of each DA, and users in each RA are not larger than m, n, and s, respectively. In the meanwhile, the line-loss of each RA is less than a constant Θ. The main parameters used in our scheme are listed in Table 2.

5.1. System Initialization. We assume the single trusted authority CC is responsible for bootstrapping the whole system. In the system initialization, use the presented security parameters to generate system parameters firstly and secondly register system entities in CC.

(i) Generating Process of System Parameters

(1) Based on the Bilinear Pairing technique, given the security parameter κ, after running $\mathscr{G}en(\kappa)$, CC can generate parameters $(q, P, \mathbb{G}_1, \mathbb{G}_2, e)$.

(2) Again, given the security parameter ξ, according to the Paillier cryptosystem, after choosing two large prime numbers p and q, whose length are both ξ, then CC can calculate the public keys $(N = pq, g)$ and the private keys (λ, μ).

(3) CC chooses a secure cryptographic hash function $H : \{0,1\}^* \longrightarrow \mathbb{G}_1$.

(4) CC also chooses two common factors Υ_1 and Υ_2 as line-loss calculation parameters randomly, which must meet the requirements that $\Upsilon_1 > \Theta, \Upsilon_2 > \Upsilon_1^n \cdot \Theta$.

(5) At last, CC publishes the system parameters $\{q, P, \mathbb{G}_1, \mathbb{G}_2, e, N, g, H, \Upsilon_1, \Upsilon_2\}$ as public keys and keeps the master key $\{\lambda, \mu\}$ secret.

(ii) Registering Process of System Entities

(1) All RAGWs send registration requests to CC; as shown in Figure 2, each RAGW belongs to a RA; the RA is in a DA. According to its actual business and administrative situation, CC sets a unique number i from the sequence set $\{1, 2, \ldots, m\}$ and another unique number j from another sequence set $\{1, 2, \ldots, n\}$ to the RAGW. The notation RAGW_{ij}

represents the RAGW which belongs to the R_{ij} (the jth RA of the ith DA).

(2) Each RAGW_{ij} ($i \in \{1,2,\ldots,m\}$, $j \in \{1,2,\ldots,n\}$) randomly chooses $x_{ij} \in \mathbb{Z}_q^*$ as its private key and calculates $Y_{ij} = x_{ij}P$ as its public key.

(3) All users in the A_{ij} send registration requests to RAGW_{ij}; as shown in Figure 3, according to the management situation, RAGW_{ij}, respectively, set a unique number k from the sequence set $\{1,2,\ldots,s\}$ to the user, as the section of system model mentioned; the notation U_{ijk} means the kth user of the jth RA in the ith DA.

(4) Each U_{ijk} ($i \in \{1,2,\ldots,m\}$, $j \in \{1,2,\ldots,n\}$, $k \in \{1,2,\ldots,s\}$) randomly chooses $x_{ijk} \in \mathbb{Z}_q^*$ as its private key, and calculates $Y_{ijk} = x_{ijk}P$ as its public key.

(5) All the VTs, which equipped with smart meters located in all RAs, also send requests to their corresponding RAGWs. Without loss of generality, considering the R_{ij}, RAGW_{ij} sets the two numbers i, j to the VTs; for the convenience of discussion and without loss of correctness, we utilize the notation $U_{ij\tau}$ for expressing the VT equipped in the R_{ij}.

(6) The VT in jth RA of the ith DA, $U_{ij\tau}$ ($i \in \{1,2,\ldots,m\}$, $j \in \{1,2,\ldots,n\}$) randomly chooses $x_{ij\tau} \in \mathbb{Z}_q^*$ as its private key and calculates $Y_{ij\tau} = x_{ij\tau}P$ as its public key.

In our system, after the two initialization processes above, we assumed the numbers of DAs, RAs in the D_i and users in the R_{ij} are $\mathcal{N}^{DA}$, $\mathcal{N}_i^{RA}$, and $\mathcal{N}_{ij}^{U}$, respectively.

5.2. Report Generation. Each U_{ijk} and $\text{VT}(U_{ij\tau})$ ($i \in \{1,2,\ldots,m\}$, $j \in \{1,2,\ldots,n\}$, $k \in \{1,2,\ldots,s\}$) can collect electricity consumption periodically (e.g., 15 minutes) by equipped with smart meters. After collecting nearly real-time electricity usage for different time $\mathcal{T}$, each user and each VT need to utilize the public keys published by CC to encrypt the data firstly, make a signature of the data by using individual private keys x_{ijk}, $x_{ij\tau}$ secondly, and generate a corresponding report thirdly.

(1) U_{ijk} and $U_{ij\tau}$ encrypt collecting data of electricity consumption d_{ijk} and $d_{ij\tau}$, respectively, as follows:

$$\begin{aligned} C_{ijk} &= g^{\Upsilon_2^i \cdot \Upsilon_1^j \cdot d_{ijk}} \cdot r_{ijk}^N \bmod N^2 \\ C_{ij\tau} &= g^{\Upsilon_2^i \cdot \Upsilon_1^j \cdot d_{ij\tau}} \cdot r_{ij\tau}^N \bmod N^2 \end{aligned} \tag{3}$$

where r_{ijk} and $r_{ij\tau}$ are both chosen by users and VTs randomly from $\mathbb{Z}_N^*$, respectively (r_{ijk}, $r_{ij\tau} \in \mathbb{Z}_N^*$).

(2) U_{ijk} and $U_{ij\tau}$ use their private keys x_{ijk} and $x_{ij\tau}$, respectively, and hash function H of CC to generate individual signatures as follows:

$$\begin{aligned} \sigma_{ijk} &= x_{ijk} \cdot H\left(C_{ijk} \parallel i \parallel j \parallel k \parallel \mathcal{T}\right) \\ \sigma_{ij\tau} &= x_{ij\tau} \cdot H\left(C_{ij\tau} \parallel i \parallel j \parallel \tau \parallel \mathcal{T}\right) \end{aligned} \tag{4}$$

(3) U_{ijk} and $U_{ij\tau}$ generate reports (see (5)), respectively:

$$\begin{aligned} D_{ijk} &= C_{ijk} \parallel i \parallel j \parallel k \parallel \mathcal{T} \parallel \sigma_{ijk} \\ D_{ij\tau} &= C_{ij\tau} \parallel i \parallel j \parallel \tau \parallel \mathcal{T} \parallel \sigma_{ij\tau} \end{aligned} \tag{5}$$

After the report is generated, U_{ijk} and $U_{ij\tau}$ send D_{ijk} and $D_{ij\tau}$ to RAGW_{ij} for each $k \in \{1,2,\ldots,\mathcal{N}_{ij}^U\}$, respectively.

5.3. Verification by RAGW. In R_{ij}, after the RAGW_{ij} receives all the reports of time $\mathcal{T}$: $\{D_{ijk}, D_{ij\tau}\}$, for each $k \in \{1,2,\ldots,\mathcal{N}_{ij}^U\}$, based on the bilinear pairings, by verifying

$$\begin{gathered} e\left(P, \sigma_{ij1}\right) \stackrel{?}{=} e\left(Y_{ij1}, H\left(C_{ij1} \parallel i \parallel j \parallel 1 \parallel \mathcal{T}\right)\right) \\ e\left(P, \sigma_{ij2}\right) \stackrel{?}{=} e\left(Y_{ij2}, H\left(C_{ij2} \parallel i \parallel j \parallel 2 \parallel \mathcal{T}\right)\right) \\ \cdots \\ e\left(P, \sigma_{ijk}\right) \stackrel{?}{=} e\left(Y_{ijk}, H\left(C_{ijk} \parallel i \parallel j \parallel k \parallel \mathcal{T}\right)\right) \\ \cdots \\ e\left(P, \sigma_{ij\mathcal{N}_{ij}^U}\right) \stackrel{?}{=} e\left(Y_{ij\mathcal{N}_{ij}^U}, H\left(C_{ij\mathcal{N}_{ij}^U} \parallel i \parallel j \parallel \mathcal{N}_{ij}^U \parallel \mathcal{T}\right)\right) \\ e\left(P, \sigma_{ij\tau}\right) \stackrel{?}{=} e\left(Y_{ij\tau}, H\left(C_{ij\tau} \parallel i \parallel j \parallel \tau \parallel \mathcal{T}\right)\right) \end{gathered} \tag{6}$$

if all the equations of (6) are hold, then reports D_{ijk} and $D_{ij\tau}$ are accepted, not vice versa. Because the function e of bilinear pairing technique is time-consuming and high cost, after introducing batch verification [10, 38], we can make the verification efficiently. The batch verification in the R_{ij} performs as

$$\begin{aligned} & e\left(P, \left(\sum_{k=1}^{\mathcal{N}_{ij}^U} \sigma_{ijk}\right) + \sigma_{ij\tau}\right) \stackrel{?}{=} e\left(P, (\sum_{k=1}^{\mathcal{N}_{ij}^U} x_{ijk}\right. \\ & \quad \cdot H\left(C_{ijk} \parallel i \parallel j \parallel k \parallel \mathcal{T}\right) + x_{ij\tau} \\ & \quad \left. \cdot H\left(C_{ij\tau} \parallel i \parallel j \parallel \tau \parallel \mathcal{T}\right)\right) \\ & \stackrel{?}{=} \left(\prod_{k=1}^{\mathcal{N}_{ij}^U} e\left(P, x_{ijk} \cdot H\left(C_{ijk} \parallel i \parallel j \parallel k \parallel \mathcal{T}\right)\right)\right) \\ & \quad \cdot e\left(P, x_{ij\tau} \cdot H\left(C_{ij\tau} \parallel i \parallel j \parallel \tau \parallel \mathcal{T}\right)\right) \\ & \stackrel{?}{=} \left(\prod_{k=1}^{\mathcal{N}_{ij}^U} e\left(Y_{ijk}, H\left(C_{ijk} \parallel i \parallel j \parallel k \parallel \mathcal{T}\right)\right)\right) \\ & \quad \cdot e\left(Y_{ij\tau}, H\left(C_{ij\tau} \parallel i \parallel j \parallel \tau \parallel \mathcal{T}\right)\right) \end{aligned} \tag{7}$$

As a result, the batch verification can reduce the running times of e from $2 \cdot (\mathcal{N}_{ij}^U + 1)$ to $\mathcal{N}_{ij}^U + 2$ compared to the original verification.

5.4. *Privacy-Preserving Line-Loss Calculation.*

After RAGWs received the verified reporters of time $\mathscr{T}$, each RAGW generates line-loss of its corresponding RA in the form of cipher by privacy-preserving calculation. Once each RA's ciphertext of line-loss is generated, all of them will be sent to CC, then CC also performs calculation to convert all the whole RAs' line-loss cipher data of time $\mathscr{T}$ to a number. The steps are as follows.

5.4.1. *Processing in RAGW.*

Let the notation $LLD\_C_{ij\mathscr{T}}$ represent a value of R_{ij} at time $\mathscr{T}$ which is calculated by

$$\begin{aligned}
LLD\_C_{ij\mathscr{T}} &= C_{ij\tau} \cdot \prod_{k=1}^{\mathscr{N}_{ij}^{U}} C_{ijk}^{-1} \bmod N^2 \\
&= \frac{C_{ij\tau}}{\prod_{k=1}^{\mathscr{N}_{ij}^{U}} C_{ijk}} \bmod N^2 \\
&= \frac{g^{\Upsilon_2^i \cdot \Upsilon_1^j \cdot d_{ij\tau}} \cdot r_{ij\tau}^{N}}{g^{\Upsilon_2^i \cdot \Upsilon_1^j \cdot d_{ij1} + \cdots + \Upsilon_2^i \cdot \Upsilon_1^j \cdot d_{ij\mathscr{N}_{ij}^{U}}} \cdot r_{ij1}^{N} \cdots r_{ij\mathscr{N}_{ij}^{U}}^{N}} \bmod N^2 \\
&= g^{\Upsilon_2^i \cdot \Upsilon_1^j \cdot d_{ij\tau} - (\Upsilon_2^i \cdot \Upsilon_1^j \cdot d_{ij1} + \cdots + \Upsilon_2^i \cdot \Upsilon_1^j \cdot d_{ij\mathscr{N}_{ij}^{U}})} r_{ij\tau}^{N} \\
&\quad \cdot r_{ij1}^{(-1)N} \cdots r_{ij\mathscr{N}_{ij}^{U}}^{(-1)N} \bmod N^2 \\
&= g^{\Upsilon_2^i \cdot \Upsilon_1^j \cdot (d_{ij\tau} - (d_{ij1} + \cdots + d_{ij\mathscr{N}_{ij}^{U}}))} \\
&\quad \cdot \left(r_{ij\tau} \cdot r_{ij1}^{-1} \cdots r_{ij\mathscr{N}_{ij}^{U}}^{-1} \right)^{N} \bmod N^2
\end{aligned} \tag{8}$$

Let

$$LLD\_P_{ij\mathscr{T}} = d_{ij\tau} - \left(d_{ij1} + \cdots + d_{ij\mathscr{N}_{ij}^{U}} \right) \tag{9}$$

According to section of system model, in (8), the notation $LLD\_P_{ij\mathscr{T}}$ means the plaintext of the line-loss of the R_{ij}, so the $LLD\_C_{ij\mathscr{T}}$ shows the line-loss in R_{ij} at time $\mathscr{T}$ in the form of cipher.

After the $LLD\_C_{ij\mathscr{T}}$ is generated, the corresponding RAGW_{ij} creates its signature σ_{ij} as

$$\sigma_{ij} = x_{ij} \cdot H\left(LLD\_C_{ij\mathscr{T}} \parallel i \parallel j \parallel \mathscr{T} \right) \tag{10}$$

Utilizing the signature, the RAGW_{ij} generates a report D_{ij} as

$$D_{ij} = LLD\_C_{ij\mathscr{T}} \parallel i \parallel j \parallel \mathscr{T} \parallel \sigma_{ij} \tag{11}$$

At last, each RAGW sends its reports of different time to CC.

5.4.2. *Processing in CC.*

Similar to the verification by RAGW, after importing the batch verification, CC can verify all the $D_{ij}s$ came from RAGWs by verifying the follow equations:

$$\begin{aligned}
&e\left(P, \sum_{i=1}^{\mathscr{N}^{DA}} \sum_{j=1}^{\mathscr{N}_i^{RA}} \sigma_{ij} \right) \\
&\stackrel{?}{=} e\left(P, \sum_{i=1}^{\mathscr{N}^{DA}} \sum_{j=1}^{\mathscr{N}_i^{RA}} x_{ij} \cdot H\left(LLD\_C_{ij\mathscr{T}} \parallel i \parallel j \parallel \mathscr{T} \right) \right) \\
&\stackrel{?}{=} \prod_{i=1}^{\mathscr{N}^{DA}} \prod_{j=1}^{\mathscr{N}_i^{RA}} \mathrm{e}\left(P, x_{ij} \cdot H\left(LLD\_C_{ij\mathscr{T}} \parallel i \parallel j \parallel \mathscr{T} \right) \right) \\
&\stackrel{?}{=} \prod_{i=1}^{\mathscr{N}^{DA}} \prod_{j=1}^{\mathscr{N}_i^{RA}} e\left(Y_{ij}, H\left(LLD\_C_{ij\mathscr{T}} \parallel i \parallel j \parallel \mathscr{T} \right) \right)
\end{aligned} \tag{12}$$

After all $D_{ij}s$ of time $\mathscr{T}$ are verified by CC, CC calculates a number $LLD\_C_{\mathscr{T}}$ to represent the line-loss originated from all RAs of the system at time $\mathscr{T}$ in the form of cipher as follows:

$$\begin{aligned}
LLD\_C_{\mathscr{T}} &= \prod_{i=1}^{\mathscr{N}^{DA}} \prod_{j=1}^{\mathscr{N}_i^{RA}} LLD\_C_{ij\mathscr{T}} \bmod N^2 \\
&= \prod_{i=1}^{\mathscr{N}^{DA}} \prod_{j=1}^{\mathscr{N}_i^{RA}} \left(g^{\Upsilon_2^i \cdot \Upsilon_1^j \cdot (d_{ij\tau} - (d_{ij1} + \cdots + d_{ij\mathscr{N}_{ij}^{U}}))} \right. \\
&\quad \left. \cdot \left(r_{ij\tau} \cdot r_{ij1}^{-1} \cdots r_{ij\mathscr{N}_{ij}^{U}}^{-1} \right)^{N} \right) \bmod N^2 \\
&= g^{\sum_{i=1}^{\mathscr{N}^{DA}} \sum_{j=1}^{\mathscr{N}_i^{RA}} \Upsilon_2^i \cdot \Upsilon_1^j \cdot (d_{ij\tau} - (d_{ij1} + \cdots + d_{ij\mathscr{N}_{ij}^{U}}))} \cdot \prod_{i=1}^{\mathscr{N}^{DA}} \prod_{j=1}^{\mathscr{N}_i^{RA}} \left(r_{ij\tau} \right. \\
&\quad \left. \cdot r_{ij1}^{-1} \cdots r_{ij\mathscr{N}_{ij}^{U}}^{-1} \right)^{N} \bmod N^2 = g^{\sum_{i=1}^{\mathscr{N}^{DA}} \sum_{j=1}^{\mathscr{N}_i^{RA}} \Upsilon_2^i \cdot \Upsilon_1^j \cdot LLD\_P_{ij\mathscr{T}}} \\
&\quad \cdot \left(\prod_{i=1}^{\mathscr{N}^{DA}} \prod_{j=1}^{\mathscr{N}_i^{RA}} \left(r_{ij\tau} \cdot r_{ij1}^{-1} \cdots r_{ij\mathscr{N}_{ij}^{U}}^{-1} \right) \right)^{N} \bmod N^2 \\
&= g^{LLD\_P_{\mathscr{T}}} \cdot \left(\prod_{i=1}^{\mathscr{N}^{DA}} \prod_{j=1}^{\mathscr{N}_i^{RA}} \left(r_{ij\tau} \cdot r_{ij1}^{-1} \cdots r_{ij\mathscr{N}_{ij}^{U}}^{-1} \right) \right)^{N} \\
&\quad \cdot \bmod N^2
\end{aligned} \tag{13}$$

(Let $LLD\_P_{\mathscr{T}} = \sum_{i=1}^{\mathscr{N}^{DA}} \sum_{j=1}^{\mathscr{N}_i^{RA}} \Upsilon_2^i \cdot \Upsilon_1^j \cdot LLD\_P_{ij\mathscr{T}}$.)

At last, CC stores a tuple $< LLD\_C_{\mathscr{T}}, \mathscr{T} >$ in database to imply the line-loss of the whole RAs at the time $\mathscr{T}$ without any privacy leaks.

```
Input: $LLD\_P_{\mathcal{T}}$
Output: Set of all $LLD\_P_{ij\mathcal{T}}$s
    for each $(i \in \{1, 2, \cdots, \mathcal{N}^{DA}\},\ j \in \{1, 2, \cdots, \mathcal{N}_i^{RA}\})$
1: for $i \longleftarrow 1$ to $\mathcal{N}^{DA}$ do
2:   for $j \longleftarrow 1$ to $\mathcal{N}_i^{RA}$ do
3:     $LLD\_P_{ij\mathcal{T}} \longleftarrow \left( \dfrac{\left( \left( LLD\_P_{\mathcal{T}} / \Upsilon_2^i \right) \bmod \Upsilon_2 \right)}{\Upsilon_1^j} \right) \bmod \Upsilon_1$
4:   end for
5: end for
6: return $\begin{bmatrix} LLD\_P_{11\mathcal{T}} & \cdots & LLD\_P_{1\mathcal{N}_1^{RA}\mathcal{T}} \\ LLD\_P_{21\mathcal{T}} & \cdots & LLD\_P_{2\mathcal{N}_2^{RA}\mathcal{T}} \\ \vdots & \ddots & \vdots \\ LLD\_P_{\mathcal{N}^{DA}1\mathcal{T}} & \cdots & LLD\_P_{\mathcal{N}^{DA}\mathcal{N}_{\mathcal{N}^{DA}}^{RA}\mathcal{T}} \end{bmatrix}$
```

Algorithm 1: Generating each $LLD\_P_{ij\mathcal{T}}$ with Horner's rule.

5.5. Decryption of Line-Loss. CC also can recover any RA's line-loss of different DAs at different time $\mathcal{T}$.

$$g^{LLD\_P_{\mathcal{T}}} \cdot \left(\prod_{i=1}^{\mathcal{N}^{DA}} \prod_{j=1}^{\mathcal{N}_i^{RA}} \left(r_{ij\tau} \cdot r_{ij1}^{-1} \cdots r_{ij\mathcal{N}_{ij}^{U}}^{-1} \right) \right)^N \bmod N^2 \tag{14}$$

Obviously, (14) coming from (13) is still in the form of ciphertext which can be generated by Paillier cryptosystem:

$$g^m \cdot r^N \bmod N^2 \tag{15}$$

where

$$LLD\_P_{\mathcal{T}} = m \tag{16}$$

$$\prod_{i=1}^{\mathcal{N}^{DA}} \prod_{j=1}^{\mathcal{N}_i^{RA}} \left(r_{ij\tau} \cdot r_{ij1}^{-1} \cdots r_{ij\mathcal{N}_{ij}^{U}}^{-1} \right) = r \tag{17}$$

Therefore, after searching from database by $\mathcal{T}$, CC can get the $LLD\_C_{\mathcal{T}}$, then by using private keys (λ, μ), the $LLD\_P_{\mathcal{T}}$ can be recovered from (14). Using Algorithm 1 can extract any $LLD\_P_{ij\mathcal{T}}$ of its corresponding R_{ij} at different time $\mathcal{T}$, and also if CC just want to get the line-loss of a specified RA $R_{\alpha\beta}$, only they need to perform the line 3 of Algorithm 1 with the parameters α and β described as (18):

$$LLD\_P_{\alpha\beta\mathcal{T}} \longleftarrow \left(\frac{\left(LLD\_P_{\mathcal{T}} / \Upsilon_2^{\alpha} \right) \bmod \Upsilon_2}{\Upsilon_1^{\beta}} \right) \bmod \Upsilon_1 \tag{18}$$

6. Correctness and Security Analysis

In this section, we analyze our scheme on correctness and security aspects. Our scheme's design is correctly proved with bilinear pairing and Horner's rule, while its security is presented in privacy preserving, authentication, and integrity.

6.1. Correctness Analysis. Before getting into details about the security of our scheme, let us firstly prove our scheme's correctness based on the properties of bilinear pairing and Horner's rule.

Theorem 1. *Each RAGW and CC will only process correct (not tampered and from a legal sender) data.*

Proof. We leverage the properties of bilinear pairing mentioned in Section 4.1 to ensure the correctness of data, so all RAGWs and CC will only process correct (not tampered and from a legal sender) data after (7) and (12) are verified, respectively. □

Theorem 2. *Using Horner's rule iteratively will recover each line-loss of all RAs or any specified RA's line-loss correctly.*

Proof. The cipher in (3) contains not only the data from corresponding SM but also the information about its region information (RA and DA) by parameters Υ_1 and Υ_2. After RAGW's processing in (8), we can also record both line-loss of the RA (see (9)) and the region information in the form of cipher. Similarly, according to (13), all line-loss data and its area information both will be stored in a cipher (number). Finally, after decryption of the Paillier by using private keys, we can get (16) which is calculated by the polynomial in (13), because of the reasonable choice of the parameters Υ_1 and Υ_2 ($\Upsilon_1 > \Theta, \Upsilon_2 > \Upsilon_1^n \cdot \Theta$), then we can use Horner's rule iteratively to get the coefficients of the polynomial then recover each line-loss of all RAs or any specified RA's line-loss correctly as shown in Algorithm 1 and (18). □

6.2. Security Analysis. Particularly, based on the security requirements discussed before, considering from three aspects, which involves privacy, authentication, and integrity, this section focuses on analyzing the security properties of the proposed EPLC scheme individually.

(i) Privacy-Preserving. Since Paillier cryptosystem had been proved semantic secure against the chosen plaintext attack, based on the proposed EPLC scheme applied in the system model, the analyses are as follows.

Theorem 3. *The users' and VTs' reports are privacy preserving in our proposed scheme.*

Proof. In our scheme, the almost real-time electricity consumption data d_{ijk} and $d_{ij\tau}$ are collected by smart meters equipped at each user (U_{ijk}) and VT($U_{ij\tau}$), which are formed and encrypted as (3).

Obviously, the ciphertexts $C_{ijk} = g^{\Upsilon_2^i \cdot (\Upsilon_1^j \cdot (-d_{ijk}))} \cdot r_{ijk}^N \bmod N^2$ and $C_{ij\tau} = g^{\Upsilon_2^i \cdot (\Upsilon_1^j \cdot d_{ij\tau})} \cdot r_{ij\tau}^N \bmod N^2$ have the same form as a valid ciphertext of Paillier cryptosystem $g^m \cdot r^N \bmod N^2$, so the data d_{ijk} in C_{ijk} and $d_{ij\tau}$ in $C_{ij\tau}$ are sematic secure and privacy preserved. Specially, since each r_{ijk} and each $r_{ij\tau}$ are a random number in $\mathbb{Z}_N^*$, based on different $r_{ijk} \neq r_{i'j'k'}$, the two same data $d_{ijk} = d_{i'j'k'}$ are encrypted to different ciphertexts $C_{ijk} \neq C_{i'j'k'}$ for resisting dictionary attacks. If the collusion attacks are launched by several users, which means all of them can share all their individual information with each other, including ID, random number r (for Paillier Encryption), and the corresponding ciphertext and in conjunction with the public information of system, no one can infer any others' private information. □

Theorem 4. *The processing in RAGW and CC are also privacy preserving in our proposed scheme.*

Proof. After having received all reports $C_{ij1}, C_{ij2}, \ldots, C_{ij\mathcal{N}_{ij}^U}$, if the adversary $\mathcal{A}$ hijacks RAGW. Firstly, all the private information of the reports received from users and VTs is encrypted by Paillier cryptosystem. Secondly, all the line-loss calculations on RAGW are based on ciphertexts generated from users and VTs. Thirdly, even if there exist collusion attacks launched by several users, VTs, and RAGWs, which means all of them can share and analyze each other's and their own information, as mentioned before, the information will involve ID, random number used by Paillier cryptosystem, and public information of the system. Because of the existence of semantic security of Paillier cryptosystem, the adversary $\mathcal{A}$ cannot infer any sensitive information of users and VTs.

The processing of CC is similar to the methods of RAGW, so in the actual operation, CC also can guarantee the users privacy away from infringement of adversary $\mathcal{A}$.

To summarise, even if the adversary $\mathcal{A}$ can eavesdrop and intercept reports on communication channel, hijack RAGW, intrude into the database of CC, and steal some data about line-loss of the system, each user's sensitive data of almost real-time electricity consumption are privacy preserved in the proposed EPLC scheme as long as the system keeps the private keys under the secret protection. □

(ii) Authentication and Integrity. Based on the *CDH* problem [37] of the random oracle model, *BLS* short signature has been proved to be secure, which is tamper-resistant and guarantees each report is from its corresponding legal sender.

Theorem 5. *The authentication and data integrity of the users' and VTs' reports, LLD_C, are both guaranteed in our proposed scheme.*

Proof. In EPLC scheme, utilizing individual private keys of each entity, reports of each individual user and VT and each ciphertext $LLD\_C_{ij\mathcal{T}}$ of line-loss calculated by RAGWs are both signed by *BLS* short signature [39] before sending. Because the secure of *BLS* short signature has been proved, which is tamper-resistant and can guarantee each report is from its corresponding legal sender, so malicious behaviors of the adversary $\mathcal{A}$ such as tampering and falsifying in the system can be detected. So the authentication and the integrity of reporters from each users, VTs, and RAGWs are guaranteed by EPLC. □

7. Performance Evaluation

In this section, we analyze the computational cost of users, RAGWs, and CC in the process of our EPLC scheme. Experiments are conducted on a PC with Processor: Intel(R) Core(TM) i5-3470 CPU @ 3.20GHz; Memory (RAM): 12 GB; OS: 64-bit Windows 10 Pro; Java-version 1.8.0_181; JPBC [40] library-version 2.0.0.; 1024-bit $N(|N^2| = 2048)$ for Paillier cryptosystem and 160-bit $\mathbb{G}_1$ for symmetric pairing.

Meter's data of users and VTs are established from the random numbers generated based on the *nextInt* method of *java.util.Random* package, and we assume the numbers of DAs and RAs in each DA and users in each RA are 10, 20, and 200; the range of values for line-loss and meter's data of users are 10 to 100 and 10 to 2000, respectively. According to the requirement of our EPLC scheme, we set Υ_1 to 101, slightly more than the maxima of line-loss and Υ_2 to 1220190039947966824482749091552564190200101, slightly more than $\Upsilon_1^{\text{The number of RAs in each DA}}$ (the maxima of line-loss).

7.1. The Analysis of Users' Computational Cost. Meter's data of users and VTs are established, followed by the encryption and signature processes. From Figure 4, it can be seen that the computational cost of users and VTs belonged to different RAs of different DAs ranging from 200*ms* to 320*ms* and the distribution is affected fairly by the sequence number of DA. According to (3) and (4), the computational cost increase with the sequence number of DA is justified.

7.2. The Analysis of RAGWs' Computational Cost. When RAGW has received all reports of users and VT in corresponding RA at a certain time $\mathcal{T}$, then it performs batch verification and line-loss calculation. It can be seen from Figure 5 that the computational cost of RAGWs belonged to different RA of different DA range from 33300*ms* to 33530*ms* and are irrelevant to the sequence number of RA and DA.

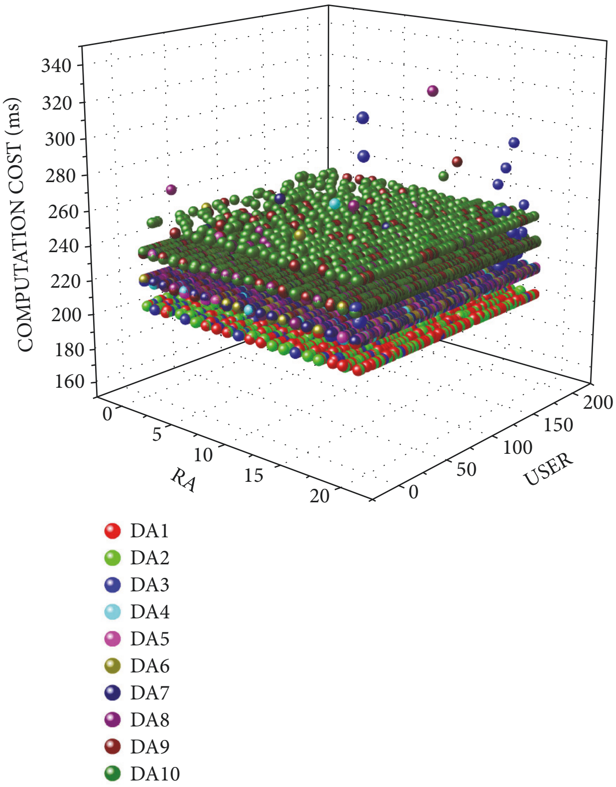


Figure 4: Computational cost of each user.

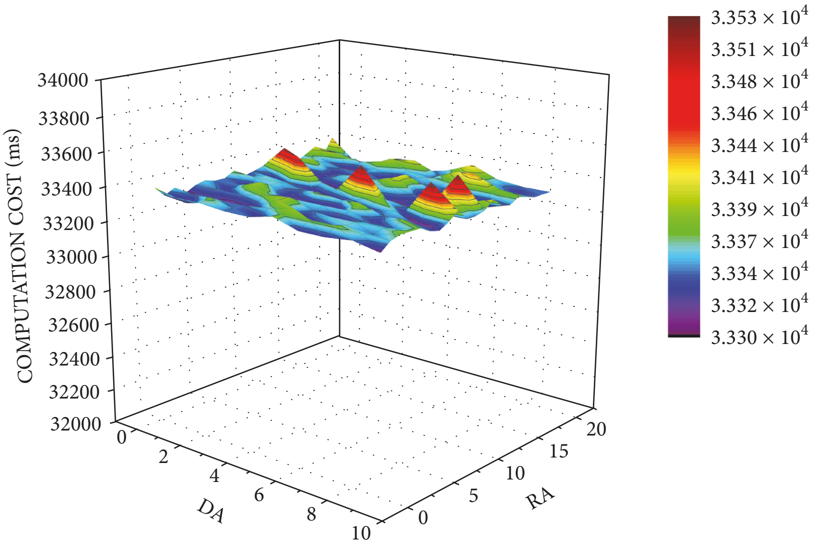


Figure 5: Computational cost of each RAGW.

7.3. The Analysis of CC's Computational Cost. Similar to RAGW, batch verification and calculation of the number $LLD\_C_{\mathcal{T}}$ will be executed when all reports belong to the certain time period $\mathcal{T}$ of RAGW received, and the computational cost of CC is 33562*ms*.

The computational cost of decryption for each line-loss of RA can be seen from Figure 6 range from 31*ms* to 47*ms*, also irrelevant to the sequence number of RA and DA.

As mentioned before, in general, the standard period of measuring and reporting by smart meters is 15 minutes [3, 4],

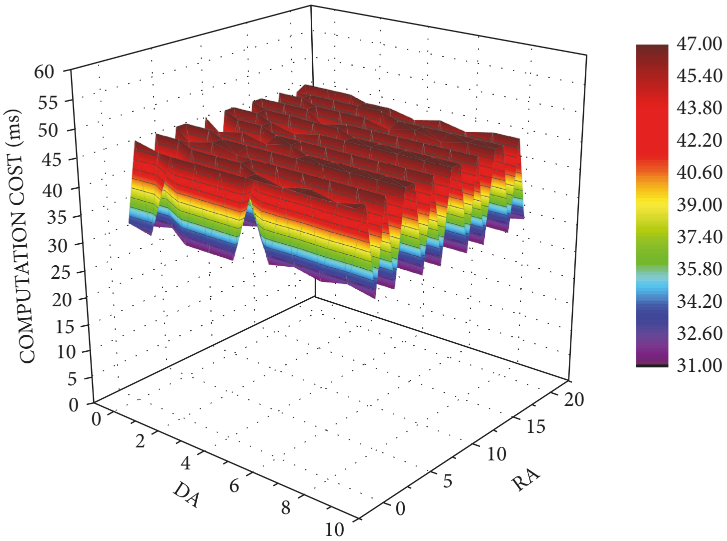


Figure 6: Computational cost of decrypting each RA line-loss.

according to the computational cost of different entities. It is obvious that our scheme achieves high efficiency.

8. Conclusion

In this paper, we have proposed an efficient privacy-preserving line-loss calculation scheme for residential areas of smart grid. The scheme can calculate the line-loss of each RA in different DA based on the Paillier cryptosystem and Horner's rule with the protection of users' privacy. For one time $\mathcal{T}$, the control center only needs to store a ciphertext $LLD\_C_{\mathcal{T}}$ to represent all values of RAs' line-loss, then can obtain each line-loss of RA in different DA when possessing the private key, and further attain finer-grained electricity regulation for smart grid. We have also demonstrated our scheme's security strength and privacy-preserving ability in the section of security analysis. And our scheme also satisfies the real-time requirement of smart grid in terms of computational cost. For the future work, line-loss of DA will be considered, and we will develop more flexible schemes to get line-loss of different areas, and then regulate electricity for RAs and DAs effectively.

Data Availability

The data used to support our findings of this study are generated by random functions of our source code.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants nos. 61572514 and 61402508); Natural Science Foundation of Hunan Province under Grant no. 2017JJ2292; Science and Technology Key Projects of Hunan Province under Grant no. 2016JC2012; Outstanding Youth Research Project of Provincial Education Department of Hunan under Grant no. 17B030; Science and Technology Planning Project of Changsha under Grants nos. K1705007, K1705018, ZD1601042, and K1705031.

References

[1] M. A. Ferrag, L. A. Maglaras, H. Janicke, J. Jiang, and L. Shu, "A systematic review of data protection and privacy preservation schemes for smart grid communications," *Sustainable Cities and Society*, vol. 38, pp. 806–835, 2018.

[2] J. Gao, Y. Xiao, J. Liu, W. Liang, and C. L. P. Chen, "A survey of communication/networking in Smart Grids," *Future Generation Computer Systems*, vol. 28, no. 2, pp. 391–404, 2012.

[3] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin, "Private memoirs of a smart meter," in *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings, BuildSys'10*, pp. 61–66, ACM, Zurich, Switzerland, November 2010.

[4] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security & Privacy*, vol. 7, no. 3, pp. 75–77, 2009.

[5] X.-H. Liu, L. Ma, Y.-S. Shi, Y.-J. Wang, and J.-H. Li, "Improved equivalent resistance method for low-voltage distribution line loss calculation," *Advanced Materials Research*, vol. 1008-1009, pp. 417–420, 2014.

[6] G. C. Ejebe, J. G. Waight, M. Santos-Nieto, and W. F. Tinney, "Fast calculation of linear available transfer capability," *IEEE Transactions on Power Systems*, vol. 15, no. 3, pp. 1112–1116, 2000.

[7] P. A. Pegoraro, J. Tang, J. Liu, F. Ponci, A. Monti, and C. Muscas, "PMU and smart metering deployment for state estimation in active distribution grids," in *Proceedings of the 2012 IEEE International Energy Conference and Exhibition*, pp. 873–878, IEEE, Florence, Italy, September 2012.

[8] K. Samarakoon, J. Wu, J. Ekanayake, and N. Jenkins, "Use of delayed smart meter measurements for distribution state

estimation," in *Proceedings of the 2011 IEEE Power and Energy Society General Meeting*, pp. 1–6, IEEE, Detroit, MI, USA, July 2011.

[9] S. Zeadally, A.-S. K. Pathan, C. Alcaraz, and M. Badra, "Towards privacy protection in smart grid," *Wireless Personal Communications*, vol. 73, no. 1, pp. 23–50, 2013.

[10] 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. 1621–1632, 2012.

[11] F. D. Garcia and B. Jacobs, "Privacy-friendly energy-metering via homomorphic encryption," in *Proceedings of the 6th International Workshop on Security and Trust Management*, pp. 226–238, Springer, Athens, Greece, 2010.

[12] S. M. Chu, M. Gong, D. S. Li, J. C. Yan, and W. P. Zhang, "Privacy-preserving smart metering," 2018, uS Patent App. 15/249,564.

[13] A. Rial and G. Danezis, "Privacy-preserving smart metering," in *Proceedings of the 10th annual ACM workshop on Privacy in the electronic society*, pp. 49–60, ACM, Chicago, IL, USA, October 2011.

[14] G. Kalogridis, R. Cepeda, S. Z. Denic, T. Lewis, and C. Efthymiou, "ElecPrivacy: Evaluating the privacy protection of electricity management algorithms," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 750–758, 2011.

[15] U. Greveler, B. Justus, and D. Loehr, "Multimedia content identification through smart meter power usage profiles," *Computers, Privacy and Data Protection*, vol. 1, p. 10, 2012.

[16] D. He, S. Chan, Y. Zhang, M. Guizani, C. Chen, and J. Bu, "An enhanced public key infrastructure to secure smart grid wireless communication networks," *IEEE Network*, vol. 28, no. 1, pp. 10–16, 2014.

[17] M. Jawurek, M. Johns, and F. Kerschbaum, "Plug-in privacy for smart metering billin," in *International Symposium on Privacy Enhancing Technologies Symposium*, vol. 6794 of *Lecture Notes in Computer Science*, pp. 192–210, Springer, Berlin, Germany, 2011.

[18] X. Lin, R. Lu, and X. S. Shen, "MDPA: multidimensional privacy-preserving aggregation scheme for wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 10, no. 6, pp. 843–856, 2010.

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

[20] Y. Sun, L. Lampe, and V. W. S. Wong, "Smart Meter Privacy: Exploiting the Potential of Household Energy Storage Units," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 69–78, 2018.

[21] G. Karopoulos, C. Ntantogian, and C. Xenakis, "MASKER: Masking for privacy-preserving aggregation in the smart grid ecosystem," *Computers & Security*, vol. 73, pp. 307–325, 2018.

[22] C. Dwork, "Differential privacy: a survey of results," in *Proceedings of the International Conference on Theory and Applications of Models of Computation*, pp. 1–19, Springer, Xi'an, China, 2008.

[23] D. Boneh, E.-J. Goh, and K. Nissim, "Evaluating 2-DNF formulas on ciphertexts," in *Theory of Cryptography*, vol. 3378 of *Lecture Notes in Computer Science*, pp. 325–341, Springer, Berlin, Germany, 2005.

[24] H. Bao and R. Lu, "A new differentially private data aggregation with fault tolerance for smart grid communications," *IEEE Internet of Things Journal*, vol. 2, no. 3, pp. 248–258, 2015.

[25] H. Shen, M. Zhang, and J. Shen, "Efficient privacy-preserving cube-data aggregation scheme for smart grids," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 6, pp. 1369–1381, 2017.

[26] S. Li, K. Xue, Q. Yang, and P. Hong, "Ppma: Privacy-preserving multi-subset aggregation in smart grid," *IEEE Transactions on Industrial Informatics*, pp. 1–10, 2017.

[27] R. Lu, K. Heung, A. H. Lashkari, and A. A. Ghorbani, "A lightweight privacy-preserving data aggregation scheme for fog computing-enhanced IoT," *IEEE Access*, vol. 5, pp. 3302–3312, 2017.

[28] L. Chen, R. Lu, and Z. Cao, "PDAFT: a privacy-preserving data aggregation scheme with fault tolerance for smart grid communications," *Peer-to-Peer Networking and Applications*, vol. 8, no. 6, pp. 1122–1132, 2015.

[29] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612-613, 1979.

[30] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 223–238, Springer, Prague, Czech Republic, 1999.

[31] J. Ni, K. Zhang, X. Lin, and X. S. Shen, "EDAT: Efficient data aggregation without TTP for privacy-assured smart metering," in *Proceedings of the 2016 IEEE International Conference on Communications, ICC 2016*, pp. 1–6, IEEE, Kuala Lumpur, Malaysia, May 2016.

[32] S. Tonyali, K. Akkaya, N. Saputro, A. S. Uluagac, and M. Nojoumian, "Privacy-preserving protocols for secure and reliable data aggregation in IoT-enabled Smart Metering systems," *Future Generation Computer Systems*, vol. 78, pp. 547–557, 2018.

[33] W. Han and Y. Xiao, "IP2DM: integrated privacy-preserving data management architecture for smart grid V2G networks," *Wireless Communications and Mobile Computing*, vol. 16, no. 17, pp. 2956–2974, 2016.

[34] M. Naehrig, K. Lauter, and V. Vaikuntanathan, "Can homomorphic encryption be practical?" in *in Proceedings of the 3rd ACM workshop on Cloud computing security workshop*, pp. 113–124, ACM, Chicago, Ill, USA, 2011.

[35] D. Boneh and M. Franklin, "Identity-based encryption from the weil pairing," in *Proceedings of the Annual international cryptology conference*, pp. 213–229, Springer, Santa Barbara, CA, USA, 2001.

[36] L. Anany, "Transform-and-conquer," in *Introduction to The Design and Analysis of Algorithms*, pp. 225–228, Addison-Wesley Longman Publishing Co., Inc., Boston, Mass, USA, 3rd edition, 2002.

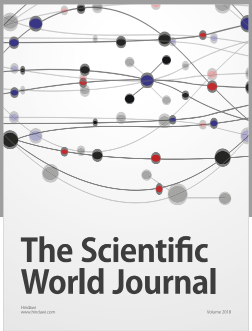
[37] M. Bellare and P. Rogaway, "Random oracles are practical: a paradigm for designing efficient protocols," in *Proceedings of the 1st ACM Conference on Computer and Communications Security*, pp. 62–73, ACM, Fairfax, VA, USA, November 1993.

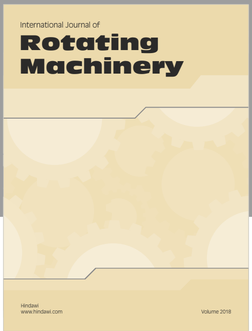
[38] C.-I. Fan, S.-Y. Huang, and Y.-L. Lai, "Privacy-enhanced data aggregation scheme against internal attackers in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 666–675, 2014.

[39] D. Boneh, B. Lynn, and H. Shacham, "Short signatures from the Weil pairing," *Journal of Cryptology*, vol. 17, no. 4, pp. 297–319, 2004.

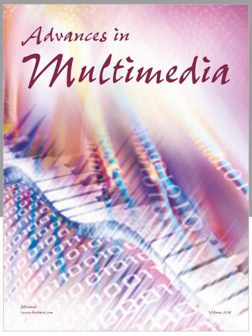
[40] A. de Caro and V. Iovino, "jPBC: Java pairing based cryptography," in *Proceedings of the 16th IEEE Symposium on Computers and Communications*, pp. 850–855, IEEE, Kerkyra, Greece, July 2011.

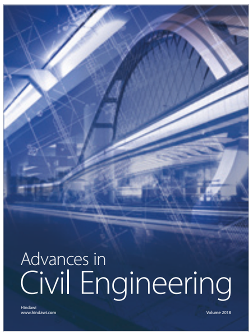
Journal of
Engineering

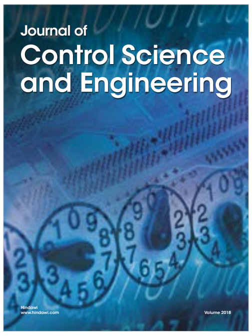

The Scientific
World Journal


International Journal of
Rotating
Machinery

Journal of
Sensors

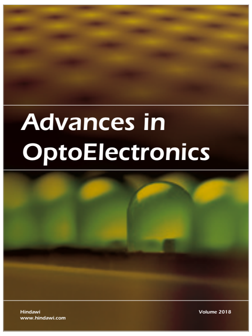

Advances in
Multimedia


Advances in
Civil Engineering


Journal of
Control Science
and Engineering

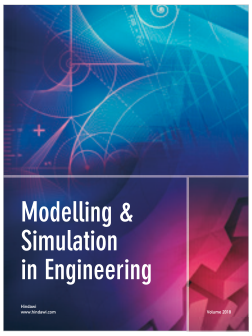
Journal of
Robotics

Journal of
Electrical and Computer
Engineering

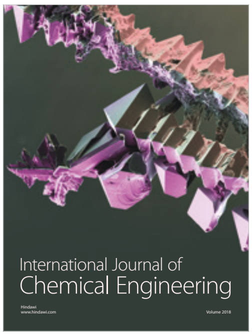

Advances in
OptoElectronics

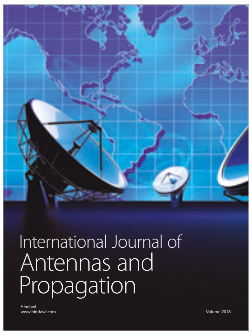
VLSI Design

International Journal of
Navigation and
Observation

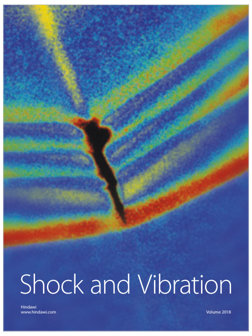

Modelling &
Simulation
in Engineering

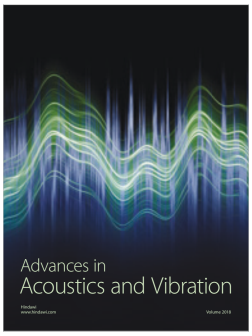
International Journal of
Aerospace
Engineering


International Journal of
Chemical Engineering


International Journal of
Antennas and
Propagation

Active and Passive
Electronic Components


Shock and Vibration


Advances in
Acoustics and Vibration