

Research Article

A Privacy-Aware Electricity Consumption Data Collection Model Based on Group Blind Signature

Fengyin Li ¹, Xiao Li ¹, Peiyu Liu ², Xueqing Sun ¹, Siqi Yu ¹ and Junrong Ge ¹

¹School of Computer Science, Qufu Normal University, Rizhao 276826, China

²School of Information Science and Engineering, Shandong Normal University, Jinan 250014, China

Correspondence should be addressed to Peiyu Liu; lpyu1960@126.com

Received 7 October 2021; Revised 14 November 2021; Accepted 27 April 2022; Published 19 May 2022

Academic Editor: Andrea Michienzi

Copyright © 2022 Fengyin Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Blockchain gives a new method for distributed data ledgering. The smart grid obtains efficient two-way data transmission and information control. It effectively monitors and regulates the grid by collecting real-time electricity consumption data of users. However, online data collection brings privacy leakage. To solve the problem of privacy leakage in the electricity data collection in the smart grid, a privacy-aware electricity data collection model is proposed. Firstly, we propose a new group blind signature scheme by introducing the blind feature into the identity-based encryption method. Secondly, by applying the proposed group blind signature scheme to the electricity data collection process, we propose a privacy-aware electricity data collection model. The proposed model ensures the conditional anonymity and traceability of user identity and the privacy protection and unforgeability of electricity consumption data.

1. Introduction

Blockchain technology originated from Satoshi Nakamoto's paper published in 2008. Blockchain, as a distributed shared ledger and database, in which records are copied and shared among its members, has the characteristics of decentralization, immutability, whole-process traces, openness, and transparency. Blockchain can store large decentralized data with better performance, availability, and scalability. Information leakage and low efficiency of blockchain are key issues that need to be addressed. A smart grid [1] is a new type of grid that combines traditional power grids with communication and information control technologies. It implements the two-way flow of the management information and power between the users and the power service provider. The architecture is shown in Figure 1. The smart grid is composed of four entities: power plant (PP), control center (CC), smart substation (SS), and smart meter (SM). There is a control center, several smart substations, and smart meters in a certain area, and the number of SSs is far less than SMs. Each SS is responsible for delivering power to users in a user area and collecting user electricity

consumption data. The SM submits the user's electricity consumption data to CC by SS. CC analyses users' electricity consumption data and arranges PP to generate power. The power arrives at SS in the form of high voltage through the high voltage transmission line. SS transforms high voltage power into low voltage power. Then, SS transmits power to a certain user area through the power distribution line.

Privacy protection issues are crucial in various systems, which are related to the reliability and security of the system. Chen et al. [2] proposed a visible routing approach PSSPR to achieve the source location privacy protection in WSNs. Li et al. [3] proposed a strong forward secure ring signature scheme based on RSA and introduced the ring signature into the privacy-aware PKI model, which achieves the privacy protection and user anonymity. Chen et al. [4] proposed a dynamic multi-key FHE scheme based on the LWE assumption in the public key setting. Otherwise, as a new biometric authentication technology, gait recognition [5–7] has attracted more and more researchers' attention in recent years. Some cloud computing-related works [8, 9] also help with privacy protection and data storage work greatly. In the smart grid, frequent information exchanges between SS and

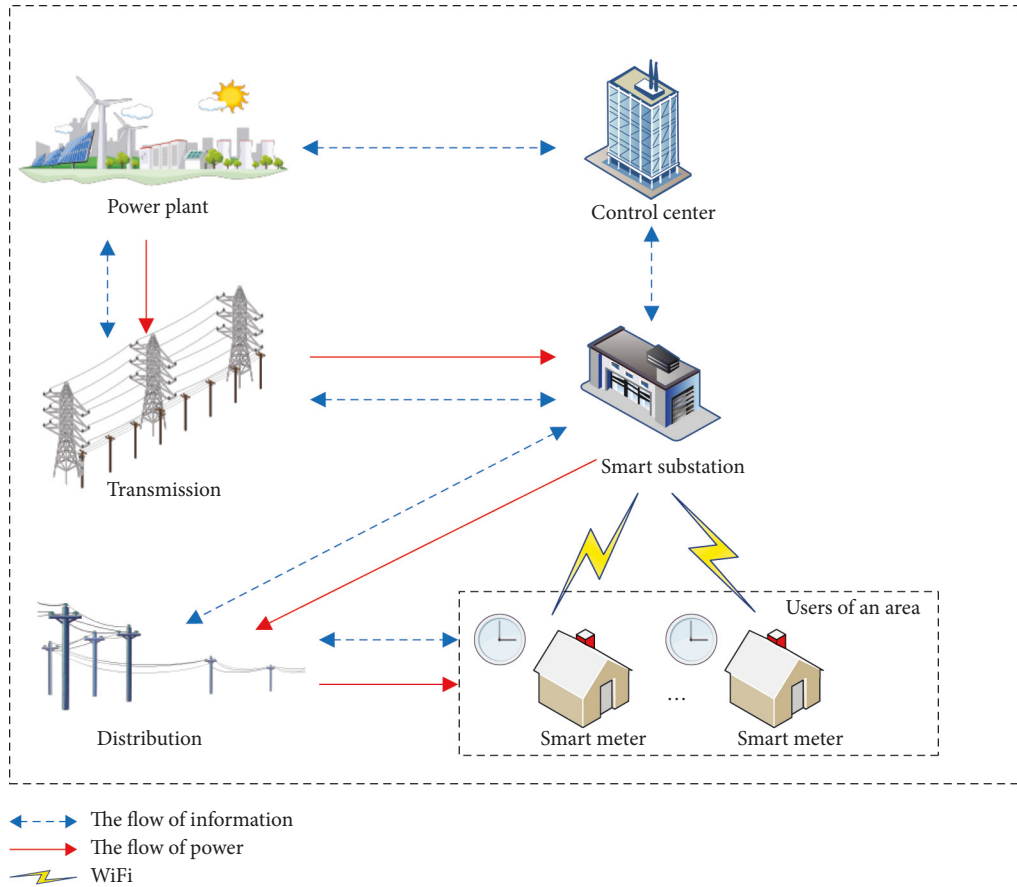


FIGURE 1: Architecture of smart grid.

SM bring privacy leakages [10, 11]. The adversary knows the user's daily schedule by eavesdropping on the electricity consumption data between SS and SM. Therefore, privacy protection in the smart grid receives more attention [12–15]. Zhao et al. [16] proposed a smart and practical privacy-preserving data aggregation scheme with smart pricing and packing method. Zhang et al. [17] proposed a blind signature-aided privacy-preserving power request scheme for a smart grid. The scheme protects the user's daily schedule. However, when the signature is invalid, CC cannot obtain the identity of the signer. The fine-grained requirements of the CC for electricity consumption data cannot be met. Some data aggregation schemes [18, 19] have been proposed in recent years. It is necessary to propose a method to implement user conditional anonymity and signer's traceability. In addition, CC obtains fine-grained electricity consumption data and verifies the integrity of the data.

Group blind signature technology provides a new way for us to achieve conditional anonymity and privacy protection for users in the smart grid. Group blind signature integrates the characteristics of group signature and blind signature at the same time. It allows the legal group member to anonymously generate signatures on behalf of the group. After the signing activity, the signer uses the group public key to verify the validity of the signature like others. However, he cannot know the signed time and who signed the signature. Due to the high anonymity and the traceability

of the group blind signature, more and more new practical schemes [20–22] have been proposed by domestic and foreign scholars. The group signature is applied in the electronic voting system [23], electronic cash system [24], intelligent transportation [25], and other fields to ensure system security. In recent years, the research of combining the group blind signature with quantum cryptography and lattice cryptography is also very popular [26–31].

In this study, we apply a new identity-based group blind signature to the privacy-aware electricity consumption data collection model. The model achieves user conditional anonymity and privacy protection. The contributions of this study are listed as follows:

- (1) By modifying the member-managing method, a new identity-based group blind signature scheme is proposed. The proposed group blind signature need not save the public keys of group members, which reduce the storage pressure of the system. The scheme effectively revokes the group members without updating the key of group manager and other group members.
- (2) Based on the proposed group blind signature scheme, a privacy-aware electricity consumption data collection model is proposed. Group blind signature assures the privacy of the electricity consumption data. In addition, we implement the user's

anonymous authentication and homomorphic verification tags, which ensure the privacy protection of user identity and the integrity of consumption data.

The organization of this study is as follows. Section 2 shows the preliminaries of this study. In Section 3, we propose a new group blind signature scheme and give its unforgeability proof. Then, we propose a privacy-aware electricity consumption data collection model based on group blind signature in Section 4. Section 5 shows the security and performance analysis of the data collection model. We conclude this study in Section 6.

2. Preliminaries

2.1. Elliptic Curve Discrete Logarithm Problem. The elliptic curve discrete logarithm problem (ECDLP) is that considering a point Q of prime order q on the elliptic curve E , if P is a possible point on E . It is difficult to find an $s \in Z_q^*$, which satisfied the equation $P = s \cdot Q$.

2.2. Group Blind Signature. A. Lysyanskaya and Z. Ramzan combined group signature and blind signature for the first time in 1998 to design the first group blind signature scheme-Lys98 scheme [32]. They used this scheme to construct an online and anonymous electronic cash system. The entities in the scheme usually contain the group manager, the group member, and the external user.

2.3. Homomorphic Tag. Homomorphism refers to mapping from one algebraic structure to another algebraic structure, and the anterior and posterior structure remains unchanged. The homomorphic tag is the tag with the property of homomorphism. Therefore, the tag of any two blocks of data can be computed from the sum of the tags of these two blocks. At the same time, when using the homomorphism tag to verify the integrity of the data, the verification can be completed only by verifying a part of the data block.

3. New Group Blind Signature Scheme

By introducing the blind feature into identity-based digital signature [33], this study proposes a new group blind signature scheme using the bilinear pairing mapping on the elliptic curve. The identity-based feature of the proposed scheme ensures that the signature system does not need to store the public key of group members. This feature reduces the storage overload of the system. In the new scheme, group members are effectively revoked without changing the key of the group manager and other group members. Security analysis indicates that the new scheme is reliable.

3.1. Group Blind Signature Scheme. Bilinear pairing is used to implement the identity-based group blind signature scheme. G_1 is an elliptic curve additive cyclic group whose order is a prime number q , and G_2 is a multiplicative cyclic group whose order is q . Meanwhile, a bilinear mapping is

$e: G_1 \times G_1 \longrightarrow G_2$. In the effective time, the discrete logarithm problem cannot be calculated whether in G_1 or G_2 .

3.1.1. System Initialization. The group manager selects generator $P \in G_1$ and three one-way hash functions: $H_1: \{0, 1\}^* \longrightarrow G_1$, $H_2: \{0, 1\}^* \longrightarrow Z_q^*$, and $H_3: G_1 \longrightarrow Z_q^*$. He chooses a random number $s \in Z_q^*$ as the private key and calculates $P_G = s \cdot P$ as his public key. Then, he initializes the group bulletin board $E = 1$ and the corresponding time T . The group manager releases system public parameters $\{G_1, G_2, P, P_G, H_1, H_2, H_3\}$ and announces the product of $H_2(ASID_i)$ on the group bulletin board, denoted as $E = \prod_i H_2(ASID_i)$.

3.1.2. Group Member Joining. A new member U_i joins this group. He first submits his real identity SID_i to the group manager. After the group manager verifies the validity of the identity, an anonymous identity code $ASID_i$ is generated. The member's public key Q_{ASID_i} and private key D_{ASID_i} are as follows:

$$Q_{ASID_i} = H_1(ASID_i), D_{ASID_i} = s \cdot Q_{ASID_i}. \quad (1)$$

The group manager saves $\langle SID_i, ASID_i \rangle$ in his database. Then, he sends the group member's private key D_{ASID_i} and anonymous identity code $ASID_i$ to U_i . At the same time, the group manager updates $E = H_2(ASID_i) \cdot E = H_2(ASID_i) \cdot \prod_i H_2(ASID_i)$. in the group bulletin board.

3.1.3. Group Member Revocation. The group manager updates the time T and E to revoke the member U_j . CC calculates $E: E = H_2(ASID_j)^{-1} \cdot E = H_2(ASID_j)^{-1} \cdot \prod_i H_2(ASID_i)$, where $j \in [0, i]$. At the same time, the group manager updates T to the present time. In this way, the group manager performs a multiplication operation to update E without changing the group public key and the group member key.

3.1.4. Group Blind Signature. For a received message, the group member signs it on behalf of the group. For instance, the signature steps of the group member $ASID_i$ are as follows:

- (1) A requester wants to acquire the signature of message m . He first chooses a random number $t_1 \in Z_q^*$ and calculates $m' = t_1 H_2(m)$. Then, he transmits m' to $ASID_i$.
- (2) After receiving m' , $ASID_i$ chooses a random number $k \in Z_q^*$ and calculates $R_1 = kP$, $S_1 = k^{-1}m'P$, and $S_2 = k^{-1}D_{ASID_i}$. Then, he sends the blind signature $\sigma' = (R_1, S_1, S_2, t)$ to requester, where t is the signature time.
- (3) requester chooses a random number $t_2 \in Z_q^*$ and calculates the signature $\sigma = (R, S, t)$ of message m as follows:

$$\begin{aligned}
R &= t_2 R_1 = t_2 k P, \\
S &= t_2^{-1} (t_1^{-1} S_1 + H_3(R) S_2), \\
&= t_2^{-1} \left(t_1^{-1} k^{-1} m P + H_3(R) k^{-1} D_{ASID_i} \right), \\
&= t_2^{-1} (t_1^{-1} k^{-1} t_1 H_2(m) P + H_3(R) k^{-1} D_{ASID_i}), \\
&= t_2^{-1} (k^{-1} (H_2(m) P + H_3(R) D_{ASID_i})).
\end{aligned} \tag{2}$$

3.1.5. Signature Verification. The validity verification of the signature $\sigma = (R, S, t)$ is divided into two steps. Firstly, the verifier selects the corresponding E based on the comparison between time t and T and verifies whether $H_2(ASID_i)$ is divisible by E . If $H_2(ASID_i)$ is not divisible, the signature is invalid. Otherwise, the signer is a member of the group. Then, the verifier uses the group public key P_G by comparing $e(R, S)$ with $e(P, P)^{H_2(m)} \cdot e(P_G, Q_{ASID_i})^{H_3(R)}$ to verify the validity of the signature. If the equation holds, σ is a validity signature. Otherwise, σ is invalid.

The verification process is as follows:

$$\begin{aligned}
e(R, S) &= e(t_2 k P, t_2^{-1} (k^{-1} (H_2(m) P + H_3(R) D_{ASID_i}))), \\
&= e(P, H_2(m) P + H_3(R) D_{ASID_i}), \\
&= e(P, P)^{H_2(m)} \cdot e(P, D_{ASID_i})^{H_3(R)}, \\
&= e(P, P)^{H_2(m)} \cdot e(P, s \cdot Q_{ASID_i})^{H_3(R)}, \\
&= e(P, P)^{H_2(m)} \cdot e(P_G, Q_{ASID_i})^{H_3(R)}.
\end{aligned} \tag{3}$$

3.2. Security Analysis. The group blind signature scheme proposed in this study satisfies unforgeability, anonymity, traceability, and revocability.

3.2.1. Unforgeability

Theorem 1. *If the ECDLP question is hard, under the existential unforgeability against chosen message attack (EU-CMA) model, the group blind signature scheme is existentially unforgeable.*

Proof. We assume that \mathcal{A} is an adversary authorized by a malicious user and able to forge group blind signatures. \mathcal{C} is a challenger who uses the adversary's ability to solve the ECDLP. However, this is contrary to the assumption of ECDLP, so the group blind signature scheme is secure. The group blind signature algorithm is modelled as a signing oracle, and the game is depicted as follows:

Setup: challenger \mathcal{C} performs the setup algorithm to generate system parameter and transmits it to \mathcal{A} . The system parameter includes $\{G_1, G_2, P, P_G, H_1, H_2, H_3\}$. \mathcal{C} randomly chooses an integer $i^* \in [1, l]$, where l denotes the maximum times of private key queries. Then, \mathcal{C} randomly chooses $s \in Z_q^*$ as the private key,

where the private key s' is equivalent to s . \mathcal{C} computes the public key $P_G = s \cdot P$.

Hash Queries: \mathcal{A} chooses the identity $ASID_i$ and sends to \mathcal{C} . \mathcal{C} calculates the hash value $Q_{ASID_i} = H_1(ASID_i)$ and sends it to \mathcal{A} .

Private Key Queries: \mathcal{A} makes the sign private key queries in this stage. \mathcal{C} maintains a list of legal signers. When \mathcal{A} queries signer's private key by sending the anonymous identity code $ASID_i$ to \mathcal{C} , \mathcal{C} checks the list of legal signers. When $i = i^*$, abort. When $i \neq i^*$, if $(i, ASID_i, Q_{ASID_i}, D_{ASID_i})$ exists, \mathcal{C} returns (Q_{ASID_i}, D_{ASID_i}) directly to \mathcal{A} . Otherwise, \mathcal{C} returns $Q_{ASID_i} = H_1(ASID_i)$ and $D_{ASID_i} = s' \cdot Q_{ASID_i}$ to \mathcal{A} and adds $(i, ASID_i, Q_{ASID_i}, D_{ASID_i})$ to the list of legal signers.

Sign Queries: adversary conducts signature queries at this stage. \mathcal{C} prepares a signature list to record all queries and responses. The list is empty at the beginning, and the format is (m, R_i, S_{1i}, S_{2i}) . \mathcal{A} selects the identity $ASID_i$ and message m , requesting the blind signature from \mathcal{C} . When \mathcal{A} queries the signature of $(m, ASID_i, Q_{ASID_i})$, if $i = i^*$, abort. Otherwise, \mathcal{C} randomly chooses $k \in Z_q^*$ and returns $\sigma'_i = (R_i, S_{1i}, S_{2i})$ to \mathcal{A} , where $R_i = kP$, $S_{1i} = k^{-1}mP$, and $S_{2i} = k^{-1}D_{ASID_i}$. Then, \mathcal{C} adds (m, R_i, S_{1i}, S_{2i}) to the signature list.

Outputs: \mathcal{A} finally outputs a valid forged signature $\sigma^* = (R_{i^*}, S_{1i^*}, S_{2i^*})$ of $ASID_{i^*}$ about the message m^* . In addition, \mathcal{A} cannot solve the ECDLP problem, so \mathcal{A} cannot get s' from $\sigma'_i = (R_i, S_{1i}, S_{2i})$. However, according to assumption \mathcal{A} gets the signature σ^* of message m^* . Therefore, \mathcal{C} obtains the solution s' of ECDLP according to the signature σ^* and the question previously queried.

Finally, \mathcal{A} solves the ECDLP assumption, but ECDLP is a difficult problem that cannot be calculated. Hence, under the difficulty assumption of ECDLP, the proposed group blind signature is existential unforgeability. \square

3.2.2. Anonymity. The correspondence $\langle SID_i, ASID_i \rangle$ between a group member's real identity SID_i and his anonymous identity code $ASID_i$ is only known by the group manager. Any other group members and external users cannot obtain it. The group member uses the anonymous identity to sign the message submitted by external users. No one obtains the real identity of the signer except the group manager, which implements the anonymity of the signer.

3.2.3. Traceability. The group member must submit his real identity SID_i to the group manager during the stage of group member joining. Then, he receives the anonymous identity code $ASID_i$ and the private key D_{ASID_i} . In this way, he becomes a legal group member and has the ability to sign messages. As long as the group member wants to correctly sign, he must use the anonymous identity code and private key distributed by the group manager. Therefore, the group manager has the ability to trace the real identity of the signer

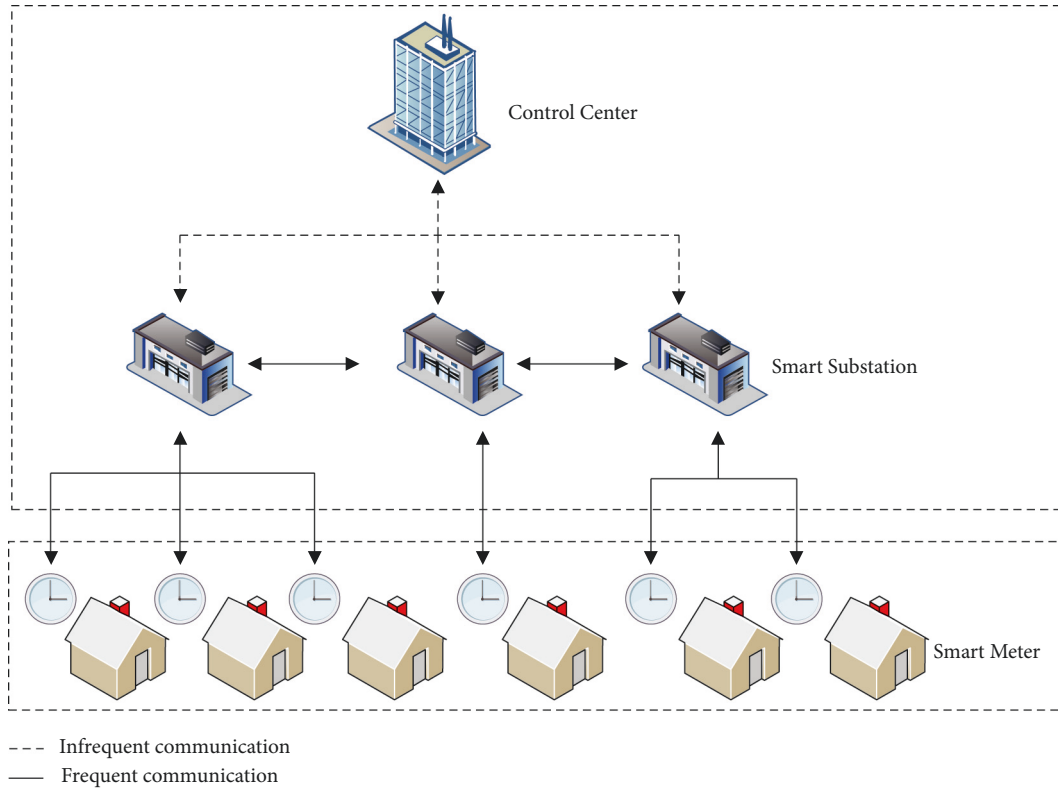


FIGURE 2: System model.

using $\langle SID_i, ASID_i \rangle$ stored in his database to achieve traceability.

3.2.4. Revocability. If a group member signs invalidly multiple times, he is identified as a malicious member. In this case, he revoked the group membership by the member revocation algorithm. Then, he loses the ability to sign on behalf of the group. Therefore, the proposed scheme has the revocability of group members.

4. A Privacy-Aware Electricity Consumption Data Collection Model Based on Group Blind Signature

By introducing the proposed group blind signature scheme into electricity consumption data collection, we propose a privacy-aware electricity consumption data collection model. The detail of the proposed model is as follows.

4.1. System Model. The system model in this study is shown in Figure 2, which involves three entities: control center (CC), smart substation (SS), and smart meter (SM). The working relationships and security requirements of the entities are as follows.

4.1.1. Control Center. CC generates system parameters, registers entities, verifies the electricity consumption data, and traces other entities conditionally. If the signature and

electricity data verification is invalid, CC traces the identity of the signer and user. CC exists in two forms, which are a fixed server located in the power plant and servers distributed in different places. CC needs to be highly credible.

4.1.2. Smart Substation. SS directly communicates with SM, verifies the user's identity, and generates the blind signature. SS does not know the user's real identity when he interacts with the user. SS needs to perform anonymous authentication on the user. SSs are fixed in certain places, generally.

4.1.3. Smart Meter. SM sends the regular electricity consumption data to CC. However, the electricity data may be tampered with within this process. Therefore, a reliable mechanism is needed to prevent the user's electricity data. SMs are installed in users' homes.

4.2. Adversary Model. The adversary model contains two main types of adversaries. One is the external adversary who is not in the data collection model. The other is the internal adversary who has the user's identity in the data collection model:

- (1) The external adversary obtains electricity consumption data by eavesdropping on the channel between SM and SS. The malicious forgery and replacement by the adversary threaten the integrity of the data.

TABLE 1: Description of notations in this study.

Notations	Descriptions
q, p	The large primes
G_1	The cyclic additive group
G_2	The cyclic multiplicative group
e	Bilinear pairing
P	A generator point
Z_q^*	Nonzero integers not larger than q
$H(\cdot)$	One-way hash function
s	Private key of group manager
P_G	Public key of group manager, where $P_G = s \cdot P$
T	The time of announcement E in the group bulletin board
E	The product of hashes of anonymous group members
SID_i	Real identity code of the group member
$ASID_i$	Anonymous identity code of the group member
D_{ASID_i}	Public key of group member, where $D_{ASID_i} = s \cdot Q_{ASID_i}$
Q_{ASID_i}	Private key of group member
m	The original message to be signed
mt	The blinded message
k	The random integer number
t_1, t_2	The blind factors
σ'	Blind signature
σ	Digital signature for m
n	Product of two large prime numbers, where $n = pq$
g	Primitive root of the modular n
a	Group public key, which is public key of RSA
b	Group private key, which is public key of RSA
x	Private key of group manager
y	Public key of group manager
$infor_i$	The information of user
gt_i	Encrypted value after user information has been hashed, where $gt_i = (H(infor_i))^x mo d n$
w_i	The random integer number
I_i	The pseudonym of user
λ	Security parameters of the electricity consumption data blocks generated by smart meters
stk	The private key of tag
ptk	The public key of tag
mx_j	The random value chosen by SM
u_j	The value needed to compute the tag is the same for each data block
tag_i	The value of tag for data block
M	The encrypted value of electricity data and the corresponding tag
TG	Multiplicative value of data block labels in a day
MG_j	The sum of the electricity data of jth dimension in a day
DG	Multiplications of bilinear pairing operation values for each data block
HS	The cumulative product of the hash value in a day

- (2) The internal adversary contains two types. One is the curious user who wants to acquire other users' electricity consumption data, but they do not tamper with any data. The other is the malicious user who tampers with his electricity consumption data.

4.3. Privacy-Aware Electricity Consumption Data Collection Model Based on Group Blind Signature. To protect the privacy of the user's identity and the electricity consumption data, an identity-based group blind signature scheme is used in the electricity consumption data collection model. CC and SSs form a group. SSs are group members, and CC is the group manager. We use the Schnorr identification protocol and the homomorphic verifiable tag mechanism to implement the anonymity of the user and the integrity verification of the electricity data. At the same time, the group blind signature mechanism ensures the traceability of the signer.

The notations used in this study are shown in Table 1.

In particular, the data collection model includes six stages: system initialization, anonymous identity authentication and data reporting, blind signature on reported electricity consumption data, data integrity verification and identity tracing, group member joining, and group member revocation. Figure 3 shows the framework of the data collection model.

4.3.1. System Initialization. In this stage, CC first generates system parameters. Then, the SS uses the real identity to apply for the group member private key and anonymous identity code. After CC verifies the identity of the SS, he distributes the anonymous identity code and the group member private key to the SS. CC saves the real identity and anonymous identity of SS in the database. SM also delivers its real information to CC and generates its own pseudonym.

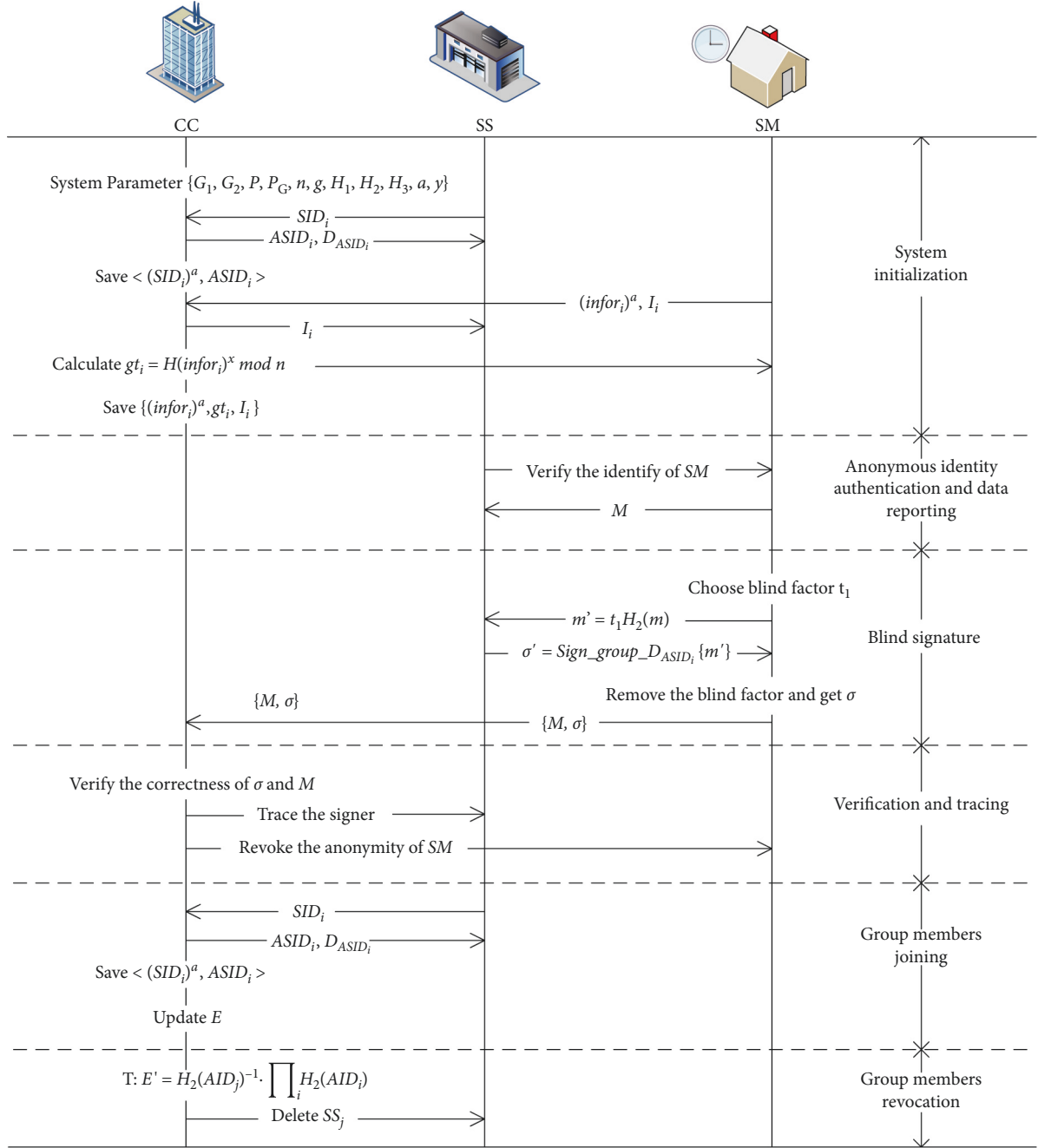


FIGURE 3: Framework of the data collection model.

CC saves the real identity and pseudonym of SM in the database. The data stored by CC, SS, and SM are, respectively, shown in Tables 2–4.

(1) Generating System Parameters.

- (i) CC computes $n = pq$, where p and q are two different large primes that meet $p|q - 1$.
- (ii) CC computes the group public key a and the private key b , where (a, b) satisfies the key pair property of RSA, namely $ab \equiv 1 \pmod{\phi(n)}$.
- (iii) CC chooses a random number $x \in [2, n - 2]$ and computes $y = g^x \bmod n$, where g is a

primitive root of the modular n . y and x are the public key and private key of the group manager, respectively.

- (iv) CC chooses generator $P \in G_1$ and three one-way hash functions: $H: \{0, 1\}^* \rightarrow \{0, 1\}^k$, $H_1: \{0, 1\}^* \rightarrow G_1$, $H_2: \{0, 1\}^* \rightarrow Z_q^*$, and $H_3: G_1 \rightarrow Z_q^*$.
- (v) CC chooses a random number $s \in Z_q^*$ as the system private key and computes $P_G = s \cdot P$ as the system public key. CC initializes the group bulletin board $E = 1$ and the corresponding time T . Then, CC releases system public

TABLE 2: Data stored by CC.

Definition	Symbol
Group public/private key	(a, b)
Group manager public/private key	(y, x)
System public/private key	(P_G, s)
The group bulletin board	E
The information of smart substation SS_i	$\langle (SID_i)^a, ASID_i \rangle (Q_{ASID_i}, D_{ASID_i})$
The information of smart meter SM_i	$(infor_i)^a H(infor_i)^x gt_i I_i$ (M, σ)

TABLE 3: Data stored by SS_i .

Definition	Symbol
Anonymous identity code	$ASID_i$
Public/private key	(Q_{ASID_i}, D_{ASID_i})
The pseudonym of SM_i	I_i
The blind signature	σ'

TABLE 4: Data stored by SM_i .

Definition	Symbol
The information of user	$infor_i$
The encrypted value of the information hash	gt_i
Random number	w_i
Pseudonym	I_i
Public/private tag key	(ptk, stk)
Electricity consumption data block and corresponding tag	(m, Tag)
The blind signature and signature of electricity consumption data	$(\sigma l, \sigma)$

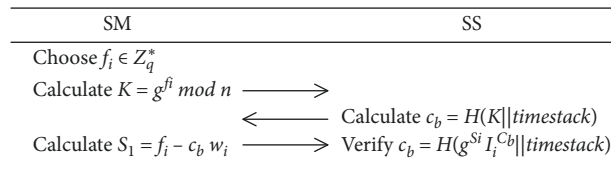


FIGURE 4: Anonymous identity authentication.

parameters $\{G_1, G_2, P, P_G, n, g, H_1, H_2, H_3, a, y\}$ and announces the time T and product of $H_2(ASID_i)$ on the group bulletin board, denoted as $E = \prod_i H_2(ASID_i)$.

(2) Registering Stage.

- (i) If SS_i wants to become a group member, he first submits real identity SID_i to CC. After CC verifies the validity of the identity, an anonymous identity code $ASID_i$ is generated. Then, CC calculates public key Q_{ASID_i} and private key D_{ASID_i} for SS_i as follows:

$$\begin{aligned} Q_{ASID_i} &= H_1(ASID_i), \\ D_{ASID_i} &= s \cdot Q_{ASID_i}. \end{aligned} \quad (4)$$

CC encrypts the real identity of the group member with the group public key a and saves $\langle (SID_i)^a, ASID_i \rangle$ in the database. Then, CC updates $E = H_2(ASID_i) \cdot E$.

- (ii) If a new user $User_i$ wants to participate in the smart grid. He first acquires $infor_i = (ID_i || \text{address} || \text{timestamp})$. Then, he encrypts his information $infor_i$ into $(infor_i)^a$ and sends it to

tag_1	tag_2	tag_3	tag_i	tag_{24}
m_1	m_2	m_3	m_i	m_{24}
m_{11}	m_{21}	m_{31}	m_{i1}	m_{241}
m_{12}	m_{22}	m_{32}	m_{i2}	m_{242}
m_{13}	m_{23}	m_{33}	m_{i3}	m_{243}
.....
m_{1j}	m_{2j}	m_{3j}	m_{ij}	m_{24j}
.....
m_{1l}	m_{3l}	m_{3l}	m_{il}	m_{24l}

FIGURE 5: Structure of electricity consumption data.

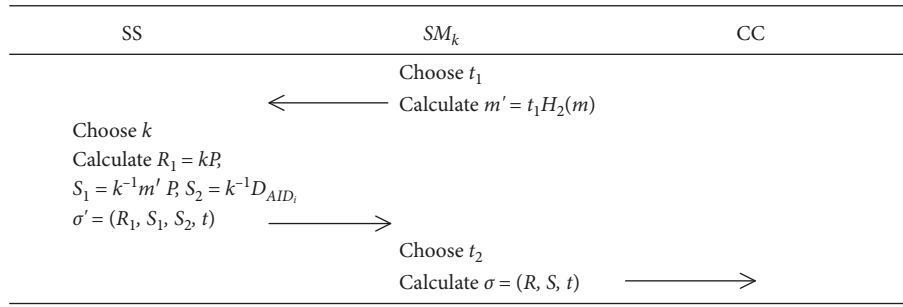


FIGURE 6: Generate the signature.

CC. CC stores $(infor_i)^a$ in his database and calculates $gt_i = (H(infor_i)^x) \bmod n$ sending it to $User_i$. The smart meters are distributed to users by CC. SM_i chooses a random number w_i to compute his pseudonym $I_i = g^{w_i} \bmod n$. SM_i sends I_i to CC.

4.3.2. Anonymous Identity Authentication and Data Reporting. In this stage, by the Schnorr identity authentication protocol, SM proves his legitimacy to SS under the condition of anonymity. Then, SM generates electricity consumption data blocks for a whole period. He calculates the data tag for each data block to ensure the integrity of data.

(1) *Anonymous Identity Authentication.* SS is not completely trusted in the model. When SM interacts with SS, the real identity of SM needs to be hidden. Therefore, the Schnorr identity authentication protocol is used to verify the legitimacy of SM. The authentication process is shown in Figure 4.

(2) *Data Reporting.* SS believes in the legitimacy of SM by anonymous identity authentication. Then, SM sends the encrypted electricity consumption data to SS. We take the example of $User_k$ encrypting and reporting electricity consumption data in one day. The whole day's data are m .

- (i) The data blocks generated in a day are restricted by the security parameter λ . We set the security

parameter λ to 24, and SM generates 24 data blocks in one day. The structure of data blocks generated in one day is shown in Figure 5. Each data block m_i represents one hour of electricity consumption data and has a corresponding tag tag_i . l -Dimensional attribute values are contained in each data block.

- (ii) SM_k randomly chooses the private tag key $stk \in Z_q^*$ and computes $ptk = gt_k^{stk} \bmod n$ as the public tag key.
- (iii) SM_k chooses l values $\{mx_1, mx_2, mx_3, \dots, mx_l\}$, randomly. Then, SM_k computes $u_j = gt_k^{mx_j} \bmod n$, where $j \in [1, l]$. SM_k calculates $tag_i = (H(MID \| i) \cdot \prod_{j=1}^l u_j^{m_{ij}})^{stk}$ for each data block m_i , where MID represents the data's summary and m_{ij} means the j th dimension attribute value of the i th data block. SM_k gets the tag set $Tag = \{tag_1, tag_2, tag_3, \dots, tag_i\}$, where the $i \in [1, 24/\lambda]$.
- (iv) SM_k computes $M = (m \| Tag)^a$ using the group public key a and calculates $H_2(m)$.

4.3.3. Blind Signature on Reported Electricity Consumption Data. In this stage, SM needs to get blind signature from SS. Then, SM reports the electricity consumption data and the signature to CC.

SS signs the electricity consumption data by the signature method provided in Section 3. SS sends the blind signature σ' to SM. SM removes the blind factor to get the signature σ . Then, SM sends σ and M to CC together. The signature generation process is shown in Figure 6.

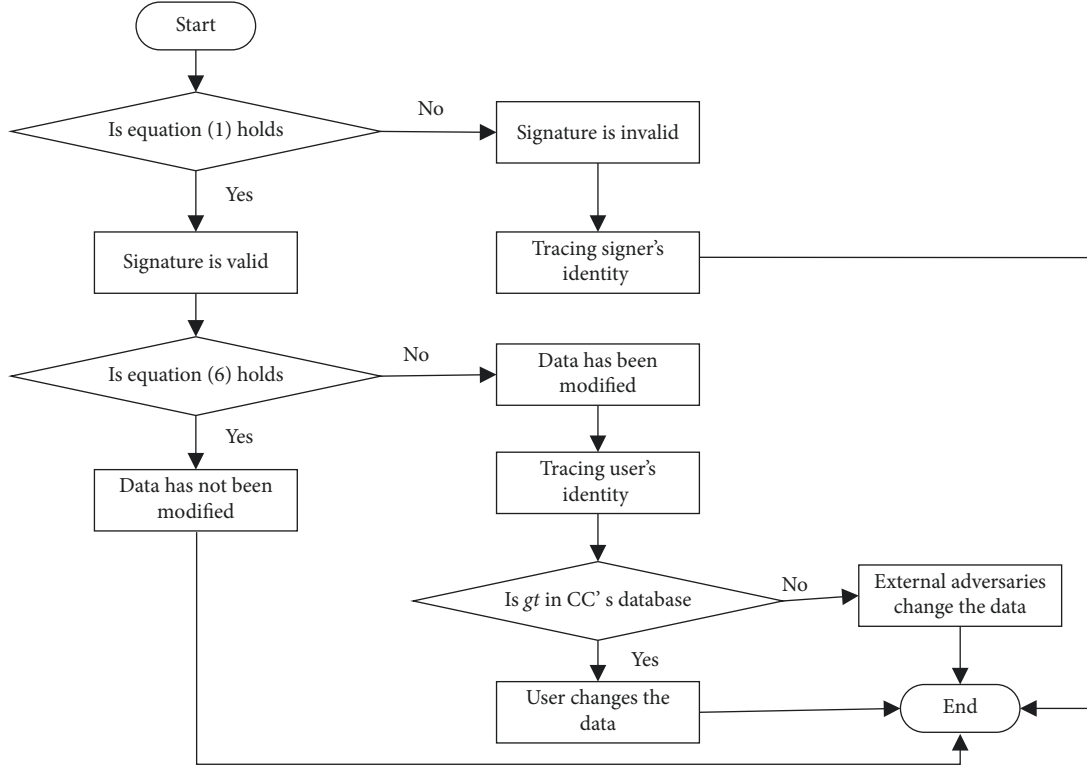


FIGURE 7: Flow chart of signature verification and identity tracing.

4.3.4. Data Integrity Verification and Identity Traceability. In this stage, CC verifies the validity of the signature and the integrity of data. Firstly, CC verifies the signature. If equation (1) holds, the signature is valid, it indicates that m has not been modified during the transmission process after being signed by SS. Otherwise, the signature is invalid. CC traces the signer's identity. Next, CC verifies the data integrity. CC uses Tag to check the integrity of the electricity consumption data. If equation (6) holds, the data are integral. Otherwise, the data have been modified. CC determines who tampered with the data. It is possible that the user or adversary has tampered with m before transmission. Therefore, CC obtains gt_k corresponding to M . CC compares gt_k with gt_i calculated using the user's real information stored in his database. If gt_k is in CC's database, it indicates that the user tampers with the data. Otherwise, the adversary tampers with the data. The flow chart of signature verification and identity tracing is shown in Figure 7.

(1) *The Verification of Signature and Data.* The verifier uses the group public key P_G to verify the validity of the signature $\sigma = (R, S, t)$. σ is the signature of M by the anonymous member $ASID_i$.

- (i) After receiving the M , CC gets the electricity consumption data m by decrypting the M . Then, CC computes $H_2(m)$ and the member's signature public key $Q_{ASID_i} = H_2(ASID_i)$.

- (ii) According to the comparison between time t and T , CC selects the corresponding E and verifies whether $H_2(ASID_i)$ is divisible by E . If $H_2(ASID_i)$ is not divisible, the signature is invalid. Otherwise, the signer is a member of the group. Then, CC uses equation 1 to verify the validity of the signature. If equation 1 holds, the signature is valid. Otherwise, the signature is invalid.

$$e(R, S) = e(P, P)^{H_2(m)} \cdot e(P_G, Q_{ASID_i})^{H_3(R)}. \quad (5)$$

- (iii) If the signature is valid, CC verifies the integrity of M . CC decrypts M to get Tag , m , u_j and calculates the following equations:

$$\begin{aligned} TG &= \prod_{i=1}^{24/\lambda} tag_i, \\ MG_j &= \sum_{i=1}^{24/\lambda} m_{ij}, \\ DG &= \prod_{j=1}^l e(u_j, ptk)^{MG_j}, \\ HS &= \prod_{i=1}^{24/\lambda} h(MID \| i). \end{aligned} \quad (6)$$

- (iv) CC verifies whether equation 6 holds every 24 hours:

$$DG \cdot e(HS, ptk) = e(TG, gt_k). \quad (7)$$

- (v) If equation 6 holds, the data have not been modified by the user or the adversary before transmission. Otherwise, the data have been modified.

(2) *Tracing the Signer and the User.*

- (i) If equation (1) does not hold, CC traces the signer's identity. The $\langle (SID_i)^a, ASID_i \rangle$ is saved in CC's database during the phase of group members joining. Therefore, CC uses the group private key b to decrypt $(SID_i)^a$ to obtain SID_i .
- (ii) If equation (1) holds and equation (6) does not hold, CC traces the user's identity to know who modified the data. CC uses the group private key b to decrypt $(infor_i)^a$ stored in his database to obtain $infor_i$. Then, CC calculates gt_i with the decrypted information one by one.

$$gt_i = H(infor_i)^x \bmod n = H(ID_i || address || timestack)^x \bmod n. \quad (8)$$

Furthermore, CC compares gt_i with gt_k , which is corresponding to M to ensure the user's identity.

4.3.5. Joining of Group Members. A new member SS_i joins the group. SS_i first sends his real identity SID_i to CC through a reliable channel. CC generates an anonymous identity code $ASID_i$, the public key Q_{ASID_i} and the private key D_{ASID_i} for the new member SS_i . Then, CC updates $E = H_2(ASID_i) \cdot E = H_2(ASID_i) \cdot \prod_i H_2(ASID_i)$ in the group bulletin board.

4.3.6. Revocation of Group Members. CC updates the time T and E published on the group bulletin board to revoke the member SS_j . CC calculates the corresponding $E = H_2(ASID_j)^{-1} \cdot \prod_i H_2(ASID_i)$ at time T , where $j \in [0, i]$.

5. Security and Performance Analysis

The security and performance analysis section shows that the proposed data collection model is secure and reliable.

5.1. Security Analysis. The security of the model is mainly based on difficult problems, such as discrete logarithm problem, elliptic curve discrete logarithm problem, and integer decomposition problem. The following shows that the proposed model has the characteristics of privacy protection, anonymity, unforgeability, and traceability.

5.1.1. Privacy Protection

Theorem 2. *Due to the difficulty of the integer decomposition problem, the adversary cannot obtain the user's electricity consumption data.*

Proof. Adversary steals m when the user reports data and obtains the blind signature stage. However, the user's

electricity consumption data m are encrypted into M by the RSA encryption method. $M = (m || Tag)^a$ can be decrypted only by the group private key b . In the data collection model, only CC has the group private key b . If the adversary wants to get m , he must obtain the private key b . The possible method is that the adversary solves the factor decomposition problem and decomposes n into correct p and q . Then, the adversary obtains the group private key b . However, the factor decomposition problem cannot be solved. The privacy protection of user electricity consumption data is implemented in our proposed model. \square

5.1.2. Anonymity. Anonymity includes the anonymity of the real identity of the SS and the real identity of the user who installed the SM.

(1) *Group Member Anonymity.* Only CC knows the correspondence $\langle (SID_i)^a, ASID_i \rangle$ between the anonymous identity and the real identity of SS. In the blind signature generation stage, the SS uses the anonymous identity to sign. Therefore, CC knows the real identity of the signer by a signature.

(2) *User Identity Anonymity.*

Theorem 3. *Because the discrete logarithm problem is difficult, \mathcal{A} cannot obtain the identity of the user by the decrypted electricity consumption data m and the corresponding tag Tag from the CC's database.*

Proof. The user's identity information $infor_i$ in CC's database is encrypted to $(infor_i)^a$. The RSA encryption is secure, and the adversary cannot calculate the group private key. Therefore, the adversary cannot obtain the user's identity information by decryption. If the adversary wants to get the user's identity, he calculates gt_k from the tag $tag_i = (H(MID || i) \cdot \prod_{j=1}^l u_j^{m_{ij}})^{stk} = (H(MID || i) \cdot \prod_{j=1}^l gt_k^{m_{ij}})^{stk}$. Then, he compares gt_k with $H(infor_i)^x \bmod n$ to determine the user's identity. However, the discrete logarithm problem is difficult, and the adversary cannot calculate gt_k from tag_i . The proposed model guarantees the anonymity of the user's identity information. \square

5.1.3. Unforgeability. Unforgeability includes the unforgeability of the group blind signature and the unforgeability of the user electricity consumption data.

(1) *Unforgeability of Group Blind Signature.* According to Theorem 1, we know whether the group blind signature is unforgeable.

(2) *Unforgeability of Electricity Consumption Data.* The adversary cannot forge the electricity consumption data. We use the homomorphic verifiable tag mechanism to verify the integrity of data. By judging whether equation (6) holds, we know whether the user's electricity consumption data have been forged or not. The detail is as follows:

$$\begin{aligned}
\text{Left} &= DG \cdot e(HS, pt_k) = \prod_{j=1}^l e(u_j, pt_k)^{MG_j} \cdot e(HS, pt_k), \\
\text{Right} &= e(TG, gt_k) = e\left(\prod_{i=1}^{24/\lambda} \text{tag}_i, gt_k\right), \\
&= e\left(\left(\prod_{i=1}^{24/\lambda} (H(MID\|i) \cdot \prod_{j=1}^l u_j^{m_{ij}})\right)^{stk}, gt_k\right), \\
&= e\left(\left(\prod_{i=1}^{24/\lambda} (H(MID|i))^{stk}\right), gt_k\right) \cdot e\left(\prod_{j=1}^l \prod_{i=1}^{24/\lambda} u_j^{m_{ij}^{stk}}, gt_k\right), \\
&= e(HS, gt_k^{stk}) \cdot e\left(\prod_{j=1}^l \sum_{i=1}^{24/\lambda} m_{ij}, gt_k^{stk}\right), \\
&= e(HS, gt_k^{stk}) \cdot e\left(\prod_{j=1}^l u_j, gt_k^{stk}\right)^{\sum_{i=1}^{24/\lambda} m_{ij}}, \\
\prod_{j=1}^l e(u_j, pt_k)^{MG_j} &= e\left(\prod_{j=1}^l \sum_{i=1}^{24/\lambda} m_{ij}, gt_k^{stk}\right), \\
&= \prod_{j=1}^l e\left(u_j, gt_k^{stk}\right)^{\sum_{i=1}^{24/\lambda} m_{ij}}, \\
&= \prod_{j=1}^l e(u_j, gt_k^{stk})^{\sum_{i=1}^{24/\lambda} m_{ij}}, \\
&= e\left(\prod_{j=1}^l u_j, gt_k^{stk}\right)^{\sum_{i=1}^{24/\lambda} m_{ij}},
\end{aligned} \tag{9}$$

Left = Right.

Therefore, we know the integrity of the user's electricity consumption data by equation (6). The proposed model guarantees the unforgeability of user electricity consumption data.

5.1.4. Traceability. As shown in Section 4.3.4, CC traces the identity of the malicious signer and user under certain conditions.

If equation (1) does not hold, CC traces the identity of the signer. CC decrypts the $(SID_i)^a$ corresponding to the signer's anonymous identity code $ASID_i$ stored in his database. $(SID_i)^a = (SID_i)^{ab} = SID_i$. Then, CC obtains the signer's real identity SID_i . If equation (6) does not hold, CC

traces the user's identity. CC gets the user's registration identity information $(infor_i)^a$, which is stored in his database. CC decrypts $(infor_i)^a$ with the group private key b to obtain $infor_i$. Then, CC calculates $gt_i = H(infor_i)^x \bmod n$ of the $infor_i$ one by one. CC compares gt_i with the gt_k , which is corresponding to the M to ensure the user's identity. If gt_i is equal to gt_k , $infor_i$ is the user's real identity.

Therefore, the proposed model guarantees the traceability of the signer's identity and the user's identity.

5.2. Performance Analysis. In the performance analysis section, we analyse the calculation cost of the electricity data

collection model in four stages, including the system initialization stage, the user authentication stage, the blind signature stage, and the verification stage.

We assume to have α smart meters and β smart substations, where B stands for bilinear pairing operation, H stands for hash operation, M stands for modular multiplication operation, L stands for modular exponentiation operation, A stands for the elliptic curve addition operation, N represents the exponential operation under the multiplication group, and W stands for the elliptic curve multiplication operation. In the system initialization stage, CC computes y , gt_i , I_i , $H(infor_i)$, Q_{ASID_i} , P_G , and D_{ASID_i} . Therefore, $L = 1 + 2\alpha$, $H = \alpha + \beta$, and $W = 1 + \beta$. In the user authentication stage, the SM computes K and S_i and the SS computes c_b , so $M = \alpha$, $L = 3\beta$, and $H = 2\alpha$. Moreover, when data reporting, one SM within a day generates ptk , u_j , tag_i , $H_2(m)$, and $H(MID||i)$. Hence, the computational cost is $M = 24/\lambda$, $L = 1 + l + 24/\lambda$, and $H = 1 + 24/\lambda$. In the blind signature stage, the calculation cost of SM acquiring signatures in a day is $W = 7\alpha$, $A = \alpha$, and $H = \alpha$. In the verification stage, the computational cost of CC verifies that signatures in a day are $B = 3\alpha$, $H = 2\alpha$, and $N = 2\alpha$. The calculation cost of verifying the data within a day for one SM is $B = l + 2$, $M = 24/\lambda$, $H = 24/\lambda$, and $L = 24/\lambda$.

6. Conclusion

This study proposes a new identity-based group blind signature scheme and applies this signature scheme to the collection of user electricity consumption data in the smart grid. Then, we obtain a privacy-aware electricity consumption data collection model based on group blind signature. The model implements the conditional anonymity of user identity information and the privacy protection of consumption data in the process of collecting electricity data. In addition, when reporting electricity consumption data, the smart meter adds a tag to the data block generated every hour through the homomorphic tag mechanism. The user's electricity consumption data for a whole day correspond to a tag set. The existence of the tag ensures the integrity and verifiability of the electricity consumption data. The security and performance analysis proves that the data collection model has privacy protection, anonymity, unforgeability, and traceability. In future work, we consider combining blockchain technology with the proposed signature scheme in the smart grid scenario to protect the privacy of the user's electricity consumption data and identity information.

Data Availability

There are no data included in this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

The authors gratefully acknowledge the helpful comments and suggestions of the reviewers, which have greatly

improved the quality of the study. This study was partly funded by EU Horizon 2020 DOMINOES Project (grant number: 771066).

References

- [1] G. Dileep, "A survey on smart grid technologies and applications," *Renewable Energy*, vol. 146, no. 1016, pp. 2589–2625, 2020.
- [2] Y. Chen, J. Sun, Y. Yang, and J. Liu, "PSSPR: a source location privacy protection scheme based on sector phantom routing in WSNs," *International Journal of Intelligent Systems*, 2021.
- [3] F. Li, Zh. Liu, T. Li, H. Ju, and W. Hua, "Privacy-aware PKI model with strong forward security," *International Journal of Intelligent Systems*, 2020.
- [4] Y. Chen, S. Dong, T. Li, Y. Wang, and H. Zhou, "Dynamic multi-key FHE in asymmetric key setting from LWE," *IEEE Transactions on Information Forensics and Security*, vol. 16, 2021.
- [5] A. Zhao, J. Li, and M. Ahmed, "Spidernet: a spiderweb graph neural network for multi-view gait recognition," *Knowledge-Based Systems*, vol. 206, Article ID 106273, 2020.
- [6] A. Zhao, J. Dong, J. Li, L. Qi, and H. Zhou, "Associated spatio-temporal capsule network for gait recognition," *IEEE Transactions on Multimedia*, vol. 24, 2021.
- [7] A. Zhao, J. Li, J. Dong, L. Qi, Q. Zhang, and N. Li, "Multimodal gait recognition for neurodegenerative diseases," *IEEE Transactions on Cybernetics*, 2021.
- [8] X. Du, S. Tang, Z. Lu, J. Wet, and K. Gai, "A novel data placement strategy for data-sharing scientific workflows in heterogeneous edge-cloud computing environments," in *Proceedings of the IEEE International Conference on Web Services (ICWS)*, pp. 498–507, Beijing, China, October 2020.
- [9] X. Du, J. Xu, and W. Cai, "OPRC: an online personalized reputation calculation model in service-oriented computing environments," *IEEE Access*, vol. 7, no. 1109, Article ID 87760, 2019.
- [10] S. Zeadally, A. S. K. Pathan, C. Alcaraz, and M. Badra, "Towards privacy protection in smart grid," *Wireless Personal Communications*, vol. 73, no. 1007, pp. 23–50, 2013.
- [11] M. Z. Gunduz and R. Das, "Cyber-security on smart grid: threats and potential solutions," *Computer Networks*, vol. 169, Article ID 107094, 2020.
- [12] M. Gough, S. Santos, T. Alskaf, M. S. Javadi, and R. Castro, "Preserving privacy of smart meter data in a smart grid environment," *IEEE Transactions on Industrial Informatics*, vol. 18, 2021.
- [13] Z. Wang, Y. Liu, Z. Ma, X. Liu, and J. Ma, "Lipsg: lightweight privacy-preserving q-learning-based energy management for the iot-enabled smart grid," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 3935–3947, 2020.
- [14] G. Tsaousoglou, K. Steriotis, N. Efthymiopoulos, and P. Makris, "Truthful, practical and privacy-aware demand response in the smart grid via a distributed and optimal mechanism," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3119–3130, 2020.
- [15] S. Uludag, S. Zeadally, and M. Badra, "Techniques, taxonomy, and challenges of privacy protection in the smart grid," *Privacy in a Digital, Networked World*, Springer, Berlin, Germany, pp. 343–390, 2015.
- [16] S. Zhao, F. Li, H. Li, R. Lu, and S. Ren, "Smart and practical privacy-preserving data aggregation for fog-based smart grids," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 521–536, 2020.

- [17] W. Zhang, Z. Guo, and N. Li, "A blind signature-aided privacy-preserving power request scheme for smart grid," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 9988170, 10 pages, 2021.
- [18] S. Li, K. Xue, Q. Yang, and P. Hong, "PPMA: privacy-preserving multisubset data aggregation in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 462–471, 2017.
- [19] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart grids using homomorphic encryption," in *Proceedings of the First IEEE International Conference on Smart Grid Communications*, pp. 327–332, Gaithersburg, MD, USA, October 2010.
- [20] Y. Yuan, Q. Li, and X. Han, "Efficient identity-based group blind signature scheme," *Computer Applications and Software*, vol. 27, no. 8, pp. 41–43, 2010.
- [21] Ch. Zhao, H. Yu, and J. Li, "Universally composable group blind signature," *Application Research of Computers*, vol. 34, no. 10, pp. 3109–3111, 2017.
- [22] W. Kong, J. Shen, P. Vijayakumar, Y. Cho, and V. Chang, "A practical group blind signature scheme for privacy protection in smart grid," *Journal of Parallel and Distributed Computing*, vol. 136, pp. 29–39, 2020.
- [23] M. Kumar, S. Chand, and C. P. Katti, "A secure end-to-end verifiable internet-voting system using identity-based blind signature," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2032–2041, 2020.
- [24] J. Zhang, Y. Yang, and S. Xie, "A third-party e-payment protocol based on quantum group blind signature," *International Journal of Theoretical Physics*, vol. 56, no. 9, pp. 2981–2989, 2017.
- [25] Y. Jiang, S. Ge, and X. Shen, "AAAS: an anonymous authentication scheme based on group signature in VANETs," *IEEE Access*, vol. 8, Article ID 98986, 2020.
- [26] X. Zhang, J. Zhang, and S. Xie, "A secure quantum voting scheme based on quantum group blind signature," *International Journal of Theoretical Physics*, vol. 59, no. 3, pp. 719–729, 2020.
- [27] R. Xu, L. Huang, W. Yang, and L. He, "Quantum group blind signature scheme without entanglement," *Optics Communications*, vol. 284, no. 14, pp. 3654–3658, 2011.
- [28] G. Liu, W. Ma, H. Cao, and L. D. Lyu, "A novel quantum group proxy blind signature scheme based on five-qubit entangled state," *International Journal of Theoretical Physics*, vol. 58, no. 6, pp. 1999–2008, 2019.
- [29] H. Zhu, Y. Tan, X. Zhang, L. Zhu, and C. Zhang, "A round-optimal lattice-based blind signature scheme for cloud services," *Future Generation Computer Systems*, vol. 73, pp. 106–114, 2017.
- [30] P. Zhang, H. Jiang, Z. Zheng, P. Hu, and Q. Xu, "A new post-quantum blind signature from lattice assumptions," *IEEE Access*, vol. 6, Article ID 27251, 2018.
- [31] K. Tiliwalidi, J. Zhang, and S. Xie, "A multi-bank E-payment protocol based on quantum proxy blind signature," *International Journal of Theoretical Physics*, vol. 58, no. 10, pp. 3510–3520, 2019.
- [32] A. Lysyanskaya and Z. Ramzan, "Group blind digital signatures: a scalable solution to electronic cash," in *Proceedings of the International Conference on Financial Cryptography*, pp. 184–197, Anguilla, February 1998.
- [33] K. G. Paterson, "ID-based signatures from pairings on elliptic curves," *Electronics Letters*, vol. 38, no. 18, pp. 1025–1026, 2002.