Blockchain-Based Decentralized Privacy-Preserving Data Aggregation (BDPDA) Scheme for SmartGrid

Hongbin Fan[1], Yining Liu[2(✉)], and Zhixin Zeng[2]

[1] College of Software and Communication Engineering, XiangNan University, Chenzhou 423000, China
[2] School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, China
ynliu@guet.edu.cn

Abstract. Smart grid is the next-generation grid that combines advanced power technology and modern communication technology. Smart meters face serious security challenges such as the leakage of user privacy and the absence of trusted third parties. Blockchain provides a viable solution that can use its key technologies to solve these problems. In blockchain technology, there is no necessary need of the third party in the energy supply sector. We introduce decentralization into smart grid, and a blockchain-based data aggregation scheme is designed. Due to the transparency of data in blockchain, the privacy of users may be disclosed. Therefore, our scheme adopts Paillier cryptosystem algorithm to encrypt the user's electricity consumption data, realizing the confidentiality of electricity consumption data, which is convenient for billing and power regulation. Through performance analysis of the scheme, it shows that the scheme has better security and better functions.

Keywords: Blockchain · Decentralized · Data aggregation · Privacy-preserving · Smart grid

1 Introduction

The traditional power grid adopts centralized power generation to meet the power demand. Therefore, the security of the centralized way is too dependent on the control center or the trusted third party. The smart grid based on the physical grid system using sensor measurement technology, communication technology, control technology and computer technology [1]. The information flow between suppliers and users in smart grid is bidirectional, while the traditional power grid adopts the unidirectional centralized system. Users can control the intelligent use of household appliances and equipment at any time according to the floating situation of electricity price in different time periods. Suppliers can automatically monitor the grid, prevent power outages, optimize grid performance, etc. However, the process of smart grid power consumption data collection may lead to the leakage of user privacy information [2, 3].

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2021
Published by Springer Nature Switzerland AG 2021. All Rights Reserved
D. Wang et al. (Eds.): SPNCE 2020, LNICST 344, pp. 448–459, 2021.
https://doi.org/10.1007/978-3-030-66922-5_31

Blockchain technology has attained significant attention recently and provides a number of ways for reliable processing and storage of data in a decentralized manner. It is now recognized that trustless decentralized energy production and transfer using blockchain technology is a promising approach. This work proposes a novel blockchain-based decentralized green energy distribution system for trustless reliable energy exchanges in a smart grid. The proof of distribution problem in a decentralized environment is first formalized. Finally, a decentralized green energy distribution smart-grid case study is presented to demonstrate the utility of the system in real-life situations.

The main contributions of this paper are as follows:

1) Decentralization: the proposed scheme is reliable as it does not rely on a trusted third party or central authority and all processes are conducted in a decentralized manner through blockchain nodes.
2) Data integrity: since the previous smart meter data aggregation methods are centralized, unlike blockchain, which allows request verification through hash mechanism. BLS signature and Paillier encryption are based on bilinear pairing, which guarantees the security and integrity of message transmission.
3) Mining node selection: The smart meters select a mining node through leader election algorithm, the mining node records smart meters' data into the blockchain.

Note that the original idea has been presented in the original conference paper. Compared with the original conference paper, this paper adds an author who participated in the revision of the paper. In the current version, the system model diagram and Blockchain structure diagram have been modified, and the graph description of leader selection algorithm and the Paillier cryptosystem introduction are added to make it easier to understand. The algorithm in this paper is compared with the existing algorithm in detail, which better reflects the advantages of the algorithm in this paper.

The rest of this paper is organized as follows. Section 2, we introduce the previous work in privacy-preserving data aggregation. Section 3, Blockchain, bilinear pairing, Boneh-Lynn-Shacham Short Signature, and the Paillier cryptosystem are given. System model and design goals are introduced in Sect. 4. In Sect. 5, our scheme is described in detail. The security analysis is in Sect. 6. In Sect. 7, the performance of our scheme is evaluated. We conclude our research in Sect. 8.

2 Related Work

In order to protect the privacy of users in smart meter data aggregation, many scholars provide various schemes.

Li et al. [4] proposed a privacy-preserving multi-subset data aggregation scheme (PPMA), their scheme based on Paillier cryptosystem, which enables the aggregation of electricity consumption data of different ranges. Liu et al. [5] proposed a privacy-preserving data aggregation without any TTP. In this scheme construct a virtual aggregation area for users with a certain degree of trust to shield the data of a single user. Guan et al. [6] proposed adjust the aggregation threshold according to the energy consumption

information and time period of each specific residential area to ensure the privacy of personal data during the aggregation process, while supporting fault tolerance. Karampour et al. [7] proposed use Paillier encryption system and AV net mask to realize the aggregation of privacy protection data in smart grid can effectively protect the privacy of user data without any security channel. However, the above research methods do not consider the trusted environment.

To achieve a trusted environment, several studies used blockchain as privacy-preserving method for data aggregation. Guan et al. [8] proposed a privacy-preserving data aggregation scheme for power grid communications. The study divided users into different groups and each group has a private blockchain. The study uses multiple pseudonyms to hide users' identity. In this scheme, key management center (KMC) is used to generate multiple public and private keys for users, which does not realize decentralization.

3 Preliminaries

In this section, we briefly introduce the necessary background.

A. Blockchain

Blockchain technology was first proposed in 2008 by Satoshi Nakamoto for Bitcoin [9]. Blockchain technology has been widely used in payment, Internet of things, healthcare, finance and so on [10]. Blockchain is a decentralized distributed ledger database maintained by network-wide nodes [11], which comprising a chain of different data blocks in a chronological order. All hash data added to the block is immutable. Blockchain is a new application mode of consensus mechanism, distributed data storage, encryption algorithm and so on. Its key technologies include block structure, Merkle tree, P2P network, hash function, timestamp, asymmetric encryption mechanism, etc. [12].

B. Boneh-Lynn-Shacham Short Signature

Boneh-Lynn-Shacham (BLS) Short Signature [13] scheme is a typical bilinear pairing scheme. The scheme uses a SHA-256 hash function $H_1 : \{0, 1\}^* \rightarrow G_1$ and g is a random generator of G_1, and a bilinear map $e : G_1 \times G_1 \rightarrow G_2$. The BLS signature scheme is divided into three phases:

1) Key generation: The secret key $x \in Z_q^*$, and compute the public key $PK = x \cdot g$.
2) Signature: The plaintext $m \in G_1$, and compute the signature $\sigma = x \cdot H(m)$
3) Verification: If $e(\sigma, g) = e(H(m), PK)$, then the signature is verified. Otherwise fails.

C. Paillier cryptosystem

Paillier cryptosystem [14] is an additive homomorphic encryption system that allows computation of encrypted data. The additive homomorphic encryption property can

directly calculate the encryption of their sum from the multiplication of the encrypted values of some data, thus effectively protecting the privacy of the data. It includes the following three algorithms:

1) Key generation: Randomly select two large primes p and q, where $|p| = |q| = |\kappa|$. Then calculate $\lambda = lcm(p-1, q-1)$. Defined a function $L(v) = \frac{v-1}{N}$, where N = pq. Choose a generator $g \in Z^*_{N^2}$, and calculate $\mu = (L(g^{\lambda} \bmod N^2))^{-1} \bmod N$. The public key is (N, g), and the corresponding private key is (λ, μ).
2) Encryption: Given a message $m \in Z_N$, choose a random number $r \in Z^*_N$. $\gcd(r, N) = 1$, The ciphertext is calculated as $C = Enc(m) = g^m \cdot r^N \bmod N^2$.
3) Decryption: Given the ciphertext $C \in Z_N$, The corresponding message is decrypted as $m = Dec(C) = L(C^{\lambda} \bmod N^2) \cdot \mu \bmod N$.

4 System Model and Design Goals

A. System model

The system model of BDPDA demonstrated in Fig. 1 consists of operation center (OC) and smart meter (SM) in the residential area (RA). In our scheme, we mainly focuses on remove the control center and the trusted third party while protecting the data privacy of the user's smart meter.

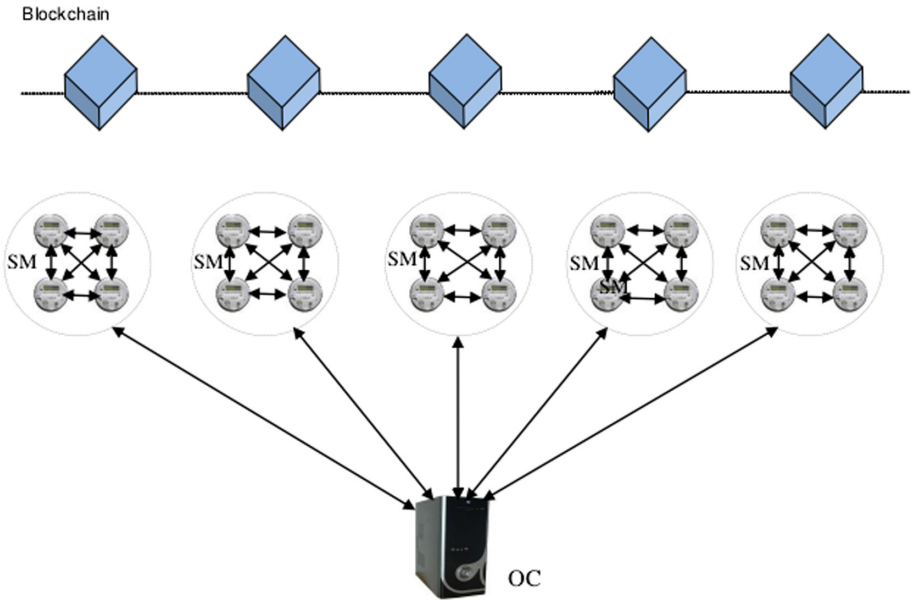

Fig. 1. System model

1) Operation center (OC)

The operation center reads the real-time power consumption data aggregated by the mining nodes of each block from the private blockchain for billing, power consumption trend analysis, adjustment of power generation plans, and dynamic pricing. To increase efficiency, each administrative district will establish its own OC.

2) Smart meter (SM)

A SM is an electricity meter for each user's site in the residential area. The smart meter regularly and simultaneously (e.g. once 15 min) collects the power consumption data of each user's household electrical equipment. Peer-to-peer (P2P) communication is used between SMs. Each residential area uses leader election algorithm to select a smart meter from the smart meters as the mining node (MN), Then each residential area constructs a block through MN.MN is responsible for generating system parameters, authenticates the legitimacy of the data transmitted by the smart meter and aggregates the encrypted data. Then, SM encrypts all kinds of collected data and uploads it to the MN after a short period of time.

B. Design goals

To solve the issues mentioned above, ensure the integrity and privacy of users' power consumption data while decentralizing and de-trusted third parties, the design goals include four aspects.

1) Privacy-preservation. Neither OC nor any other user has access to the user's data in the residential area. An external adversary cannot obtain the user's power consumption data, even if he knows the cipher text and public key. When the adversary and OC collude with each other, they can't get the power consumption data of a single user's smart meter.
2) Decentralizing. The BDPDA scheme does not need a trusted third party or central authority.

5 The Proposed Scheme

In this section, the data aggregation of decentralized smart grid based on blockchain is proposed. The notations are listed in Table 1.

A. System Initialization

OC collects electricity consumption data of L residential areas. There are n smart meters in RAj. Through leader election algorithm, selects a SM as a mining node from the n SMs in RAj, then constructs the jth block, where MNj is the root of the Merkle tree in the jth block. The state change of leader election algorithm is shown in Fig. 2. The consumption data of SMs in RAj is aggregated to MN through Merkle tree. The structure of Blockchain is shown in Fig. 3.

Table 1. Notations

Symbol	Definition
g1, g2	A generator of G
RA_j	The Jth residential area
m_i	Power consumption data of the ith smart meter in RA_j
n	Number of smart meters in the jth residential area
H1	Hash functions: H1 : {0,1}* → G
L	Number of residential areas
SM_i	Smart meter in jth residential area
MN_j	Mining node of the jth residential area
M_j	The aggregated electricity consumption data of the j-th residential areas
‖	Concatenation operation

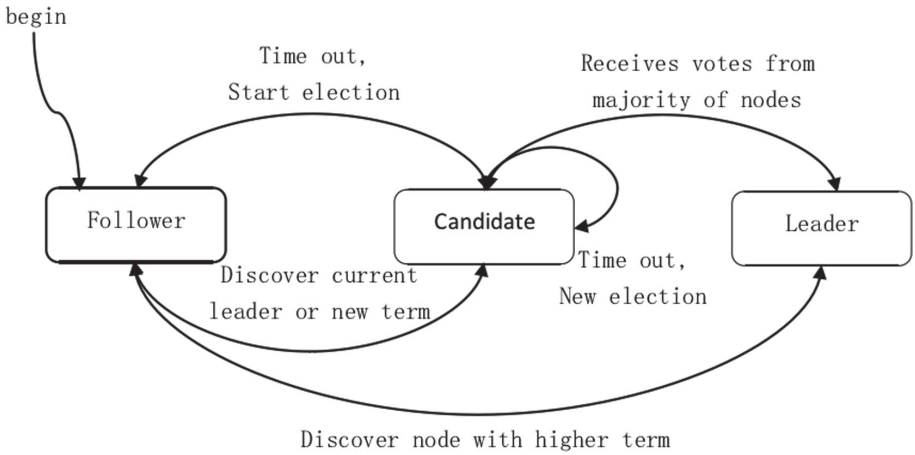

Fig. 2. State transition model of leader election algorithm

MNj runs Bilinear parameter generator $Gen(\kappa)$ to generate (q, g_1, G_1, G_2, e), and g_1 is a generator of G_1. MN calculates Paillier cryptosystem public key (N, g_2), corresponding private key (λ, μ), $g_2 \in Z^*_{N^2}$. MN choose a SHA-256 hash function H_1 and a secure cryptographic hash function H_2, where $H_1 : \{0, 1\}^* \rightarrow G_1, H_2 : \{0, 1\}^* \rightarrow \{0, 1\}^\kappa$.

MNj publishes the system public parameter $\{q, g_1, g_2, G_1, G_2, e, N, H_1\}$.

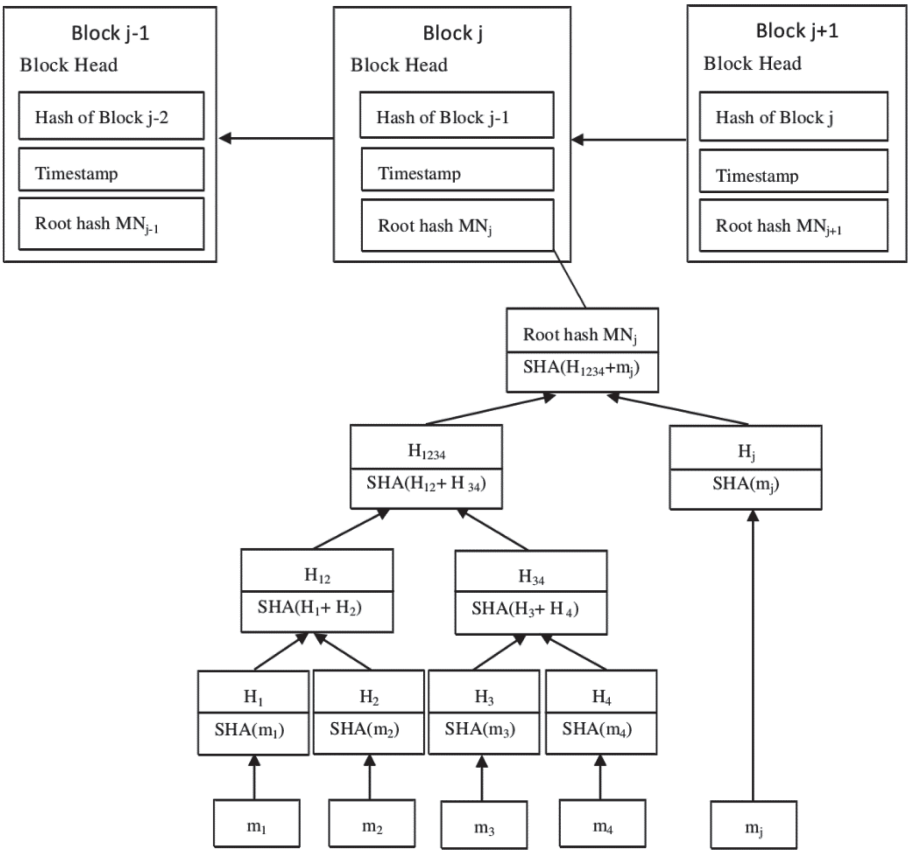

Fig. 3. Blockchain structure in our scheme

B. Ciphertext generation

1) Step 1: SM_i selects a random number $x_i \in Z_q^*$ as the private key and computes the corresponding public key $PK_i = x_i \cdot g_1$.
2) Step 2: SM_i collects electricity consumption data at timestamp T, and computes the Hash value H_2 (T), then selects a random number $r_i \in Z_N^*$ to generate ciphertext: $C_i = g_2^{m_i} \times (r_i \times H_2(T))^N \bmod N^2$.
3) Step 3: SM_i generates the BLS short signature $\sigma_i = x_i \cdot H_1(C_i \| PK_i \| Ts_i)$, Ts_i is the current timestamp to prevent replay attack.
4) Step 4: SM_i sends $C_i \| PK_i \| Ts_i \| \sigma_i$ to MN through the Merkle tree.

C. Ciphertext aggregation

1) Step 1: After receives $C_i \| PK_i \| Ts_i \| \sigma_i$, $\mathrm{MN_j}$ verifies whether $e(\sigma_i, g_1) \stackrel{?}{=} e(H_1(C_i \| PK_i \| Ts_i), PK_i)$

hold, the signature is valid and $\mathrm{MN_j}$ will accept SM_i's ciphertext. In order to make the verification more efficient, $\mathrm{MN_j}$ adopts batch verification.

2) Step 2: $\mathrm{MN_j}$ aggregates the ciphertext.

$$C = \prod_{i=1}^{n} C_i = \prod_{i=1}^{n} g_2 \cdot (r_i \cdot H_2(T))^N \bmod \mathrm{N}^2 = g_2^{\sum_{i=1}^{n} m_i} \cdot \prod_{i=1}^{n} (r_i \cdot H_2(T))^N \bmod \mathrm{N}^2$$

D. Ciphertext Decryption

$\mathrm{MN_j}$ uses the private key (λ, μ) to decrypt the aggregated ciphertext to obtain the aggregated electricity consumption data $\mathrm{M_j}$ of the j-th residential district.

$$M_j = L(C^{\lambda} \bmod \mathrm{N}^2) \cdot \mu \bmod \mathrm{N} = \sum_{i=1}^{n} m_i$$

E. Data reading

MNj generates the j+1th block, and adds the j-th block to the blockchain after the j-1 block. OC obtains the power consumption data through the public key read blockchain.

6 Security Analysis

A. Privacy-preserving

When an external attacker invades a smart meter, only the ciphertext Ci sent by a smart meter can be obtained. Even if the malicious user intercepts the ciphertext Ci, because he/she does not know the decryption key λ of the Paillier encryption algorithm, he/she cannot decrypt the ciphertext Ci to obtain the power consumption data of a single user. The power consumption data of a single smart meter is not disclosed, so as to protect the privacy of users.

B. Decentralized

In our scheme, the blockchain can be implemented without a trusted third party or central authority, the availability and reliability of data is guaranteed by MN election. No single organization can control or run SM. P2P network is adopted among smart meters to realize decentralization. The whole process does not rely on a trusted third party to make our solution more reliable and convenient.

7 Performance Evaluation

The performance of BDPDA is evaluated in this section, including the computation complexity of SM and OC, and the communication overhead.

A. Computation complexity

Compared with multiplication operation and exponentiation operation, Leader election and Hash operation is negligible. In the BDPDA scheme, the computations in the data aggregation process mainly include three phases, data encryption, batch verification and aggregation, decryption. We denote the computational cost of an exponentiation operation and a multiplication operation, by T_{exp}, T_{mul}, respectively. The computation complexities of the major entities in the system are as show in Table 2.

Table 2. Comparing computation complexity between the proposed scheme and other schemes

Scheme Ref.	Overhead SM	Overhead GW	Overhead CC	Overhead MN
[4]	$3T_{exp} + 4T_{mul}$	nT_{mul}	$T_{exp} + (4n + 3)T_{mul}$	–
[6]	$4T_{exp} + 3T_{mul}$	$3T_{exp} + (2n + 1)T_{mul}$	$3T_{exp} + 2T_{mul}$	–
[7]	$2T_{exp} + nT_{mul}$	nT_{mul}	$T_{exp} + 3 - T_{mul}$	–
BDPDA	$2T_{exp} + 4T_{mul}$	–		$T_{exp} + (n + 1)T_{mul}$

We conduct the experiments with the cpabe0.10 [15] library on a 3.0 GHz-processor and a 2 GB memory PC. As shown in Fig. 4, our scheme has advantage in computational overhead compared with PPMA, EFFECT and Karampour's schemes.

B. Communication overhead

The communication of the proposed BDPDA scheme is only SMi to MNj. Suppose that SMi generates a 2048-bit ciphertext Ci and chooses 160-bit Z_N^*. Table 3 shows the communication overhead of our scheme compared with the other three schemes with n users. It is obvious that our BDPDA scheme has a lower total communication cost than other schemes.

C. Comparison with Existing Schemes

This section describes the comparison of the proposed scheme with the existing schemes. The comparison results show that schemes [4, 6] and [7] are not based on blockchain and cannot achieve decentralization. Although scheme [8] is based on blockchain, it uses Key management center to generate public and private keys, relying on trusted

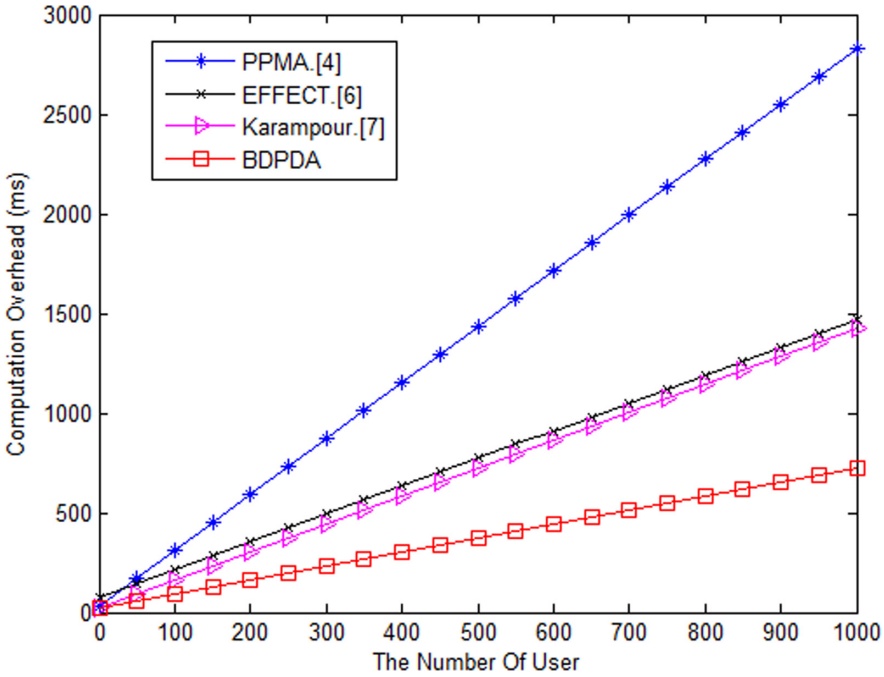

Fig. 4. Comparison of computational cost

Table 3. Comparing communication cost between the proposed scheme and other schemes

Scheme Ref.	SM-to-SM (bit)	SM-to-GW (bit)	GW-to-CC (bit)	SM-to-MN(bit)
[4]	–	2048n	2048	–
[6]	–	2048n	2048	–
[7]	n(2048(n − 1))	2048n	2048	–
BDPDA	–	–	–	2048n

third party, so decentralization cannot be achieved. Therefore, as shown in Table 4, we can see that the scheme proposed in this paper can protect user privacy while achieving decentralization.

Table 4. Comparison between proposed scheme and other related schemes

Security requirements	[4]	[6]	[7]	[8]	Our scheme
Blockchain-Based	No	No	No	Yes	Yes
Decentralization	No	No	No	No	Yes
Privacy	Yes	Yes	Yes	Yes	Yes
Non-repudiation	No	Yes	No	Yes	Yes
Confidentiality	Yes	Yes	Yes	Yes	Yes
Data Integrity	Yes	Yes	Yes	Yes	Yes
Replay attack resistance	No	Yes	Yes	Yes	Yes
Data unforgeability	No	Yes	Yes	Yes	Yes

8 Conclusion

In this paper, BDPDA scheme for smart grid is proposed, the blockchain is used to implement the decentralization of data privacy protection for the smart grid. The smart meters select a mining node through election algorithm, the mining node records smart meters' data into the blockchain. BLS signature and Paillier encryption are based on bilinear pairing, which guarantees the security and integrity of message transmission. The security analysis has proven that our mechanism meets the requirements of privacy protection and security of smart grids. The performance evaluation shows that our scheme has advantage compared with some popular data aggregation schemes in computational efficiency. As future work, we will study the combination of blockchain and other algorithms to aggregate multidimensional data and fault tolerance.

References

1. Fang, X., Misra, S., Xue, G., Yang, D.: Smart grid-the new and improved power grid: a survey. IEEE Commun. Surv. Tutor. **14**(4), 944–980 (2011)
2. Xue, K.P., Li, S.H., Hong, J.N., et al.: Two-cloud secure database for numeric-related SQL range queries with privacy preserving. IEEE Trans. Inf. Forensics Secur. **12**, 1596–1608 (2017)
3. Wu, J., Dong, M., Ota, K., Liang, L., Zhou, Z.: Securing distributed storage for social internet of things using regenerating code and Blom key agreement. Peer-to-Peer Netw. Appl. **8**, 1133–1142 (2015). https://doi.org/10.1007/s12083-014-0286-y
4. Li, S., Xue, K., Yang, Q., Hong, P.: PPMA: privacy-preserving multisubset data aggregation in smart grid. IEEE Trans. Ind. Inf. **14**, 462–471 (2018)
5. Liu, Y., Guo, W., Fan, C., Chang, L., Cheng, C.: A practical privacy-preserving data aggregation (3PDA) scheme for smart grid. IEEE Trans. Ind. Inform. **15**, 1767–1774 (2018)
6. Guan, Z., Zhang, Y., Zhu, L., Wu, L., Yu, S.: EFFECT: an efficient flexible privacy-preserving data aggregation scheme with authentication in smart grid. Sci. China Inf. Sci. **62**(3), 32103 (2019). https://doi.org/10.1007/s11432-018-9451-y

7. Karampour, A., Ashouri-Talouki, M., Ladani, B.T.: An efficient privacy-preserving data aggregation scheme in smart grid. In: 2019 27th Iranian Conference on Electrical Engineering (ICEE), pp. 1967–1971. IEEE (2019)
8. Guan, Z.T., et al.: Privacy-preserving and efficient aggregation based on blockchain for power grid communications in smart communities. IEEE Commun. Mag. **56**(7), 82–88 (2018)
9. Nakamoto, S. Bitcoin: a peer-to-peer electronic cash system (2008, consulted)
10. Crosby, M., Pattanayak, P., Verma, S., et al.: Blockchain technology: beyond bitcoin. Appl. Innov. **2**(6–10), 71 (2016)
11. Yuan, Y., Wang, F.-Y.: Parallel blockchain: concept, methods and issues. Acta Autom. Sinica **43**(10), 1703–1712 (2017)
12. Xie, Q.H.: Research on blockchain technology and financial business innovation. Financ. Dev. Res. **5**, 77–82 (2017)
13. Boneh, D., Lynn, B., Shacham, H.: Short signatures from the Weil pairing. In: Boyd, C. (ed.) ASIACRYPT 2001. LNCS, vol. 2248, pp. 514–532. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45682-1_30
14. Paillier, P.: Public-key cryptosystems based on composite degree residuosity classes. In: Stern, J. (ed.) EUROCRYPT 1999. LNCS, vol. 1592, pp. 223–238. Springer, Heidelberg (1999). https://doi.org/10.1007/3-540-48910-X_16
15. Bethencourt, J.: Advanced Crypto Software Collection: The CPABE Toolkit (2018). http://acsc.cs.utexas.edu/cpabe/