Advanced Metering Infrastructure Data Aggregation Scheme Based on Blockchain

Hongliang TIAN, Naiqian ZHENG, Yuzhi JIAN
School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, China

***Abstract*—Smart grid stands as both the cornerstone of the modern energy system and the pivotal technology for addressing energy-related challenges. Advanced Metering Infrastructure constitute a critical component within the smart grid ecosystem, providing real-time energy consumption data to power utility companies. Advanced Metering Infrastructure enables these companies to make timely and accurate decisions. Hence, the issue of data security pertaining to Advanced Metering Infrastructure assumes profound significance. Presently, Advanced Metering Infrastructure data confronts challenges associated with centralized data storage, rendering it susceptible to potential cyberattacks. Moreover, with the burgeoning number of electricity consumers, the resultant data volumes have swelled considerably. Consequently, the transmission of this data becomes intricate and its efficiency is compromised. To address these issues, this paper presents a lightweight blockchain data aggregation scheme. By integrating fog computing and cloud computing, a three-tier blockchain-based architecture is devised. Initially, digital signatures are employed to ensure the validity and integrity of user data. The innate attributes of blockchain technology are harnessed to safeguard the security of electricity energy data. Through secondary data aggregation, the privacy-sensitive user data is efficiently compressed and subsequently integrated into the blockchain, thereby mitigating the storage pressure on the blockchain and enhancing data transmission efficiency. Ultimately, through rigorous theoretical analysis and simulated experimentation, the paper demonstrates that, in comparison to existing methodologies, lightweight blockchain data aggregation scheme exhibits heightened security. Additionally, lightweight blockchain data aggregation scheme holds a competitive advantage in terms of computational and communication costs.**

***Keywords*—Smart grid; blockchain; advanced metering infrastructure; data aggregation**

I. Introduction

Smart grid epitomizes the evolution and transformation of the power and energy industry, constituting a pivotal platform for the execution of new energy strategies and the optimization of energy resource allocation. Functioning as an intelligent power network, the smart grid is constructed upon the foundation of an integrated and high-speed bidirectional communication network. Leveraging cutting-edge sensing and measurement technologies, advanced equipment methodologies, sophisticated control techniques, and state-of-the-art decision support systems, the power grid is imbued with traits of reliability, safety, cost-effectiveness, efficiency, environmental consciousness, and user security [1]. The deployment of the smart grid has notably catalyzed interaction between users and power utility companies, fostering a two-way exchange of power and data that has substantiated substantial economic and societal gains. The amalgamation of distributed generation, user-oriented energy consumption management, and remote monitoring has been actualized [2]. The realization of these outcome rests upon the bedrock of essential components like smart meters, the proliferation of which has engendered a prodigious volume of corresponding power data. The effective transmission and processing of this data stands as a vital prerequisite for the smart grid's success, as underscored by computational and communication cost considerations [3]. Moreover, the security and confidentiality of this data are of paramount importance, given its role in shaping customer electricity billing and guiding the decisions of power utility companies. Data security concerns encompass tampering, data falsification, and database attacks. In this context, blockchain technology, a decentralized data processing paradigm, emerges as a robust safeguard, with all network nodes being collectively responsible for data storage, thus ensuring comprehensive data security. In tandem with the continuous advancement of blockchain technology, numerous countries have synergistically integrated smart grid infrastructures with blockchain [4]. Within the smart grid domain, Advanced Metering Infrastructure (AMI) store substantial volumes of private user data, thereby adopting blockchain to fortify AMI data protection and storage within the blockchain, effectively mitigating the potential repercussions of data breaches [5]. Simultaneously, leveraging its distributed architecture, fog computing facilitates computations at the network's edge. In comparison to cloud computing, fog computing holds distinct advantages in processing smart meter data [6]. Notably, within smart meters, fog nodes can expedite the processing of user privacy data [7]. Given the shared architectural underpinnings of fog computing and blockchain, their integration holds great potential. Building upon prior research, this paper introduces a lightweight data aggregation schema (LDAS-BC), amalgamating fog computing, blockchain technology, and the Paillier homomorphic encryption algorithm. This schema refines the Paillier encryption system and employs a two-tier aggregation model to achieve granular data aggregation. Moreover, the scheme leverages the lightweight, one-way irreversible properties of hash algorithms to authenticate components, thereby minimizing computational and communication overheads. Overall, the primary contributions of this endeavor are as follows:

1) The introduction of cloud computing and fog computing in a three-tier architecture—comprising the user layer, fog layer, and cloud layer—attenuates performance

limitations arising from the constrained storage and computing resources of network edge devices. The intermediary fog node routinely collects user data from smart meters and significantly enhances data transmission efficiency through secondary aggregation operations.

2) The LDAS-BC scheme harnesses blockchain technology and harnesses fog nodes to formulate a fog chain. This dual-pronged approach not only permits fault tolerance for select fog nodes via consensus and master node selection algorithms but also furnishes clouds with stable and dependable data services.

3) By synergizing improved additive homomorphic encryption techniques, the schema presents a lightweight data aggregation mechanism predicated on blockchain to ensure the privacy of aggregated and transmitted data. Simultaneously, digital signatures founded on hash algorithms underwrite data integrity and validity during transmission.

The ensuing sections delineate the research structure: Section II reviews pertinent literature; Section III expounds upon the network model of the schema; Section IV introduces the master node algorithm tasked with electing a master node from the entire pool; Section V elaborates on the blockchain-rooted data aggregation mechanism. Section VI is dedicated to simulation experiments and performance analyses, culminating in a final comparison of data. Lastly, Section VII encapsulates this paper's findings and outlines potential future directions.

II. Related Work

In recent years, several privacy-conscious data aggregation methodologies have emerged, aiming to safeguard the confidentiality of transmitted data within the smart grid context. Chen et al. [8] introduced a data aggregation scheme founded on the Paillier homomorphic encryption algorithm. Nevertheless, this scheme neglected the constrained computational capabilities of smart meters and was incapable of executing pairing operations. Liang et al. [9] proposed a protocol centered on total homomorphic encryption, yet the intricate implementation of total homomorphic encryption posed a challenge. Gope et al. [10] devised a gradual data aggregation scheme that employed an aggregation tree to consolidate users' energy consumption data. The mechanism required all smart meters to partake in the aggregation process to ensure scheme accuracy. Despite this safeguard against aggregation interruption due to smart meter failure through Ping tests and third-party aggregator (TPA) involvement, the scheme incurred substantial communication overhead. Furthermore, excessive reliance on TPA engendered issues of trust and single points of failure. Singh et al. [11] proposed a privacy-ensuring data aggregation model integrating deep learning and homomorphic encryption to mitigate the adverse effects of flash memory workload on predictive model accuracy. The model facilitated secure data aggregation at low computational costs. However, it failed to account for data volume and suffered from sluggish transmission efficiency.

And in the context of blockchain. Guan et al. [12] introduced a blockchain-based methodology for privacy protection and secure, efficient data aggregation. The scheme employed pseudonyms to mask user identities and maintained data on a private blockchain. Identity authentication primarily relied on Bloom Filters within this framework. Chen et al. [13] advanced a dual blockchain-supported secure and anonymous data aggregation protocol named DA-SADA. The scheme formed a three-tier data aggregation structure via fog computing, incorporating secure and anonymous data aggregation mechanisms involving Paillier additive homomorphic encryption, aggregate signatures, and anonymous authentication, with minimal computational overhead. Faiza et al. [14] conceived PrivDA, a blockchain and homomorphic encryption-based privacy-preserving IoT data aggregation approach, enabling consumer users to create smart contracts to stipulate terms of service and requested IoT data. Cristina et al. [15] innovated a privacy-enhancing distributed security protocol leveraging blockchain and homomorphic encryption for data aggregation, employing homomorphic encryption for data encryption and blockchain smart contracts for aggregation. Bao et al. [16] advanced the BBNP paradigm, a blockchain-grounded model employing data aggregation protocols to safeguard data privacy and communication confidentiality, deploying identity authentication mechanisms for data integrity, and employing the subjective logical reputation model for consensus to address single point of failure. Zhang et al. [17] presented a potent blockchain-oriented multidimensional data aggregation framework using the Byzantine consensus mechanism to designate master nodes for data aggregation and secret sharing-based user management. Zhao et al. [18] introduced a blockchain-rooted privacy protection billing framework underpinned by the BGN encryption scheme, safeguarding users' private billing data. Notwithstanding, none of these solutions offered comprehensive data protection within the smart grid milieu. Yu et al. [19] proposed a privacy-preserving data aggregation and quality assessment protocol driven by smart contracts, storing data on the Inter Planetary File System (IPFS), deriving summary outcomes to evaluate data quality, and allocating rewards based on data quality.

In smart grid and edge computing scenarios. Lu et al. [20] introduced an edge blockchain-aided lightweight privacy-protecting data scheme dubbed EBDA within the smart grid context. Zhang et al. [21] devised LPDA-EC, a lightweight privacy-preserving data aggregation framework tailored for edge computing, ensuring data confidentiality and privacy. Fan et al. [22] introduced DPPDA, a distributed privacy-protecting data aggregation methodology for the smart grid, uniting master node algorithms, the Paillier encryption system, Boneh-Lynn-Shacham short signatures, and SHA-256 capabilities to meet security and privacy requisites for data aggregation in a fledgling grid, ensuring equitable and secure smart grid communications.

III. Smart Grid Network Model Based on Blockchain

A. Main Objectives of the System Model

Currently, substantial challenges persist in ensuring data privacy within the smart grid framework. Firstly, at the user level, the possibility of malevolent entities fabricating spurious data for transmission to fog nodes or tampering with data conveyed by smart meters poses a threat to data validity and

integrity. Consequently, data source verification becomes imperative to establish the authenticity and integrity of transmitted data. Secondly, the issue of raw data exposure remains prominent, as malicious actors could potentially gain access to such data. To mitigate this, encryption of original data is necessary, safeguarding data security from the inception of transmission through storage. Thirdly, the quandary pertains to data storage methodology. The conventional practice of storing data on fog node servers renders it susceptible to malevolent infiltration, jeopardizing power data confidentiality. To counter this, blockchain serves as a robust solution, averting single points of failure and safeguarding against server attacks by storing data in a decentralized manner. Lastly, as user numbers proliferate, so does the voluminous power consumption data. This exponential data growth necessitates data compression to curtail transmission duration, enhance efficiency, and facilitate swift decision-making at the cloud computing center. In light of these challenges, this paper introduces the LDAS-BC scheme, chiefly targeting data integrity, validity, and privacy, while also addressing the efficacy of data storage and transmission.

B. System Model

This section introduces the blockchain-centered smart grid network model, as proposed within the scope of this paper. The schematic representation of this network model is depicted in Fig. 1. The blockchain-based smart grid network model comprises three distinct strata: the user layer, the fog layer, and the cloud layer.

1) User layer: This layer predominantly encompasses an extensive array of smart meters, which, in alignment with their geographic distribution, establish connections with adjacent fog nodes. Concurrently, each smart meter dispatches encrypted data in the form of user reports to the respective area's fog node.

2) Fog layer: Comprising fog nodes endowed with computational capabilities, this layer undertakes the verification of received user reports. Once confirmed as accurate, these nodes perform primary aggregation operations on the ciphertext received from the user layer. Subsequently, an elected master node oversees secondary aggregation, culminating in the storage of the resultant aggregated data onto the blockchain, followed by data transmission to the cloud.

3) Cloud layer: The cloud layer is chiefly responsible for system initialization, data retrieval, decryption, storage, and analysis, among other functionalities. The cloud server is capable of disseminating electricity consumption reports, deduced via analysis, to the associated fog nodes. This facilitates low-latency real-time electricity consumption data queries for users.

The crux of the network model comprises three core entities: the smart meter (SM), the fog node, and the measurement data management system (MDMS). Within this intricate network framework, the Advanced Metering Infrastructure (AMI) network is methodically subdivided into 'm' distinct subregions. Each such subregion encompasses 'n' smart meters, primarily tasked with the collection of users' energy consumption data. $SM_{ij}\left(0 \le i \le n, 0 \le j \le n\right)$ is the i-th smart meter in region j, and smart meters form a user layer. In each region of the user layer, a fog node is strategically positioned for deployment. These fog nodes undertake the pivotal role of data concentration and possess the capability to aggregate the collected data. Spatially situated at the network's periphery, these fog nodes bridge the gap between the user layer and the cloud infrastructure. The measurement data management system resides within the cloud environment, equipped with the capacity to access aggregated data residing on the fog chain. Upon retrieval, this system is proficient in decryption procedures, thereby enabling analysis of the decrypted data. The outcomes of this analysis are subsequently harnessed to formulate pertinent supply strategies. Notably, the MDMS encompasses functionalities akin to cloud servers and reputable third-party entities, thereby encapsulating multifaceted roles within its purview.

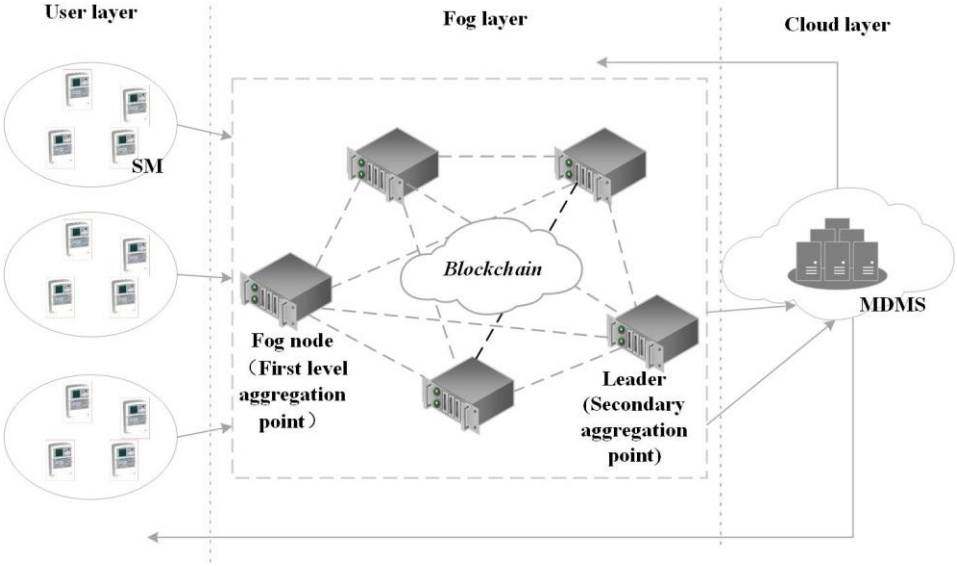

Fig. 1. Network model of LDAS-BC.

IV. Election of the Leader Node

In the realm of blockchain technology, the master node assumes a pivotal role as a fundamental component within the blockchain network. Its principal functions encompass the verification of operations executed by other nodes, the preservation of blockchain integrity, and the packaging of transactions into fresh blocks for incorporation into the blockchain. Within the context of fog nodes, the role of a primary node becomes essential in performing analogous tasks. This section sets forth a novel approach for the election of a primary node, designated to undertake the aforementioned operations.

Algorithm1 Election of the Master Node

1 When $n > 3f + 1$, $DC_i \rightarrow Follower(i \in 1,2,...,n)$, where f is the number of faulty nodes;
2 Set the term number to 0, that is $TN_{DC_i} = 0(i \in 1,2,...,n)$;
3 Set the initial vote number to 0, that is $N_v = 0$;
4 Start timing and express as $Times$;
5 Set a time threshold, that is T_{out};
6 While $Times > T_{out}$ do;
7 $Follower \rightarrow Candidate$;
8 $TN + 1$;
9 $Times$ reset to zero and restart the timer;
10 $N_v + 1$;
11 Send voting requests to other nodes and wait for responses;
12 if the replies from other nodes are received, then the cumulative number of votes N_v is calculated;
13 if $N_v > n/2 + 1$, Where n is the number of nodes, then $Candidate \rightarrow Leader$;
14 end If
15 else (Primary node has been identified);
16 $Candidate \rightarrow Follower$;
17 else
18 Repeat steps 7-11 to start a new election;
19 end If
20 end While

The master node, pivotal in linking aggregated data, subsequently transmitting this data to the cloud, is underscored by three distinct states within each fog node: candidate (sole initiator of the election), follower (participant in the voting process), and master node (sole entity authorized to modify operations). The algorithm governing the master node unfolds across three phases, as delineated below, and is encapsulated in the steps presented within Algorithm 1.

1) Preparation stage: In cases where a primary node is non-existent, the primary node election process is activated. All fog nodes are initialized with a term of 0, and the initial vote tally stands at 0.

2) Voting stage: Upon exceeding the designated time threshold, all nodes transition from follower nodes to candidate nodes, thereby heralding the commencement of the voting process.

3) End stage: Upon the vote count of a given node exceeding half of the total votes, that node ascends to the status of primary node. Subsequently, this node broadcasts its identity as the primary node. Other candidate nodes, in turn, assume the status of followers.

This method ensures the seamless selection of a primary node, vital for executing essential operations within the fog node domain.

V. Data Aggregation Mechanism Based on Blockchain

This section elucidates the LDAS-BC scheme delineated in this paper, which ingeniously amalgamates blockchain technology with Paillier homomorphic encryption technology. The scheme encompasses four integral components: system initialization, user report generation, fog chain generation, and data reading and analysis. A visual representation of these components is visually depicted in Fig. 2.

A. System Initialization

Select *k* as the system security parameter, and compute the Paillier algorithm, public key $(n = p \cdot q, g)$, private key $(\lambda = lcm(p-1, q-1), \mu)$, where p, q are large prime numbers satisfying the $|p| = |q| = |k|$ condition, let $g = N + 1$, ensure that $\mu = \left(L\left(g^{\lambda} \bmod N^2\right)\right)^{-1} \bmod N$ exists. The system randomly selects $r \in Z_N^*$ and calculates $s = r^N \bmod N^2$. Define the function $L(u) = u - 1/N$ and select the safe hash function $h: \{0,1\}^* \rightarrow \{0,1\}^l$.

B. Generation of User Reports

The energy consumption data acquired from smart meters inherently harbors users' confidential information, necessitating the encryption of such data during the collection process. For contextual clarity, this paper postulates a scenario where the private data gathered by a smart meter is transmitted to a fog node in 15-minute intervals. It's worth noting that the smart meter employs encryption mechanisms to secure the energy consumption data every 15 minutes, ensuring data integrity and privacy through digital signatures. During a designated timeframe, each fog node undertakes the aggregation of data submitted by individual smart meters.

1) SM_{ij} generates the corresponding energy consumption data, which is represented by Formula 1, including user ID_{ij}, energy consumption data d_{ij} and time stamp T_p.

$$m_{ij} = ID_{ij} \| d_{ij} \| T_p \tag{1}$$

2) Use the corresponding key (g, n, s) to calculate ciphertext C_{ij}, which can be expressed as Formula 2.

$$C_{ij} = g^{d_{ij}} \cdot s = (N+1)^{d_{ij}} \cdot s = (1 + d_{ij} n) \cdot s \tag{2}$$

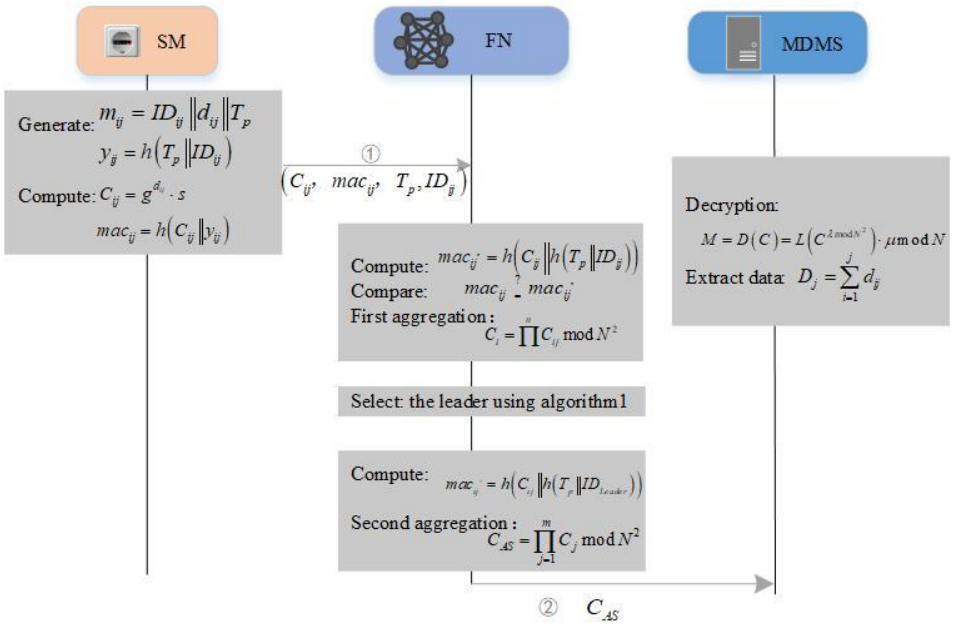

Fig. 2. LDAS-BC working flowchart.

3) After data encryption is completed, SM_{ij} uses hash function h to calculate the digital signature, which can be expressed as Formula 3, where $y_{ij} = h(T_p \| ID_{ij})$.

$$mac_{ij} = h(C_{ij} \| y_{ij}) \tag{3}$$

4) SM_{ij} sends the user report $(C_{ij}, mac_{ij}, T_p, ID_{ij})$ to the fog node.

C. Generation of Fog Chain

This section unfolds across five primary stages, each uniquely handling data in distinct ways. The algorithmic representation of state transitions for each fog node is illustrated in Algorithm 1. During the initial three stages, all fog nodes actively engage, while subsequent to the selection of the primary node, only the fog node designated as the primary node partakes in the final two stages.

1) Data verification: Upon the receipt of a user report, the fog node conducts digital signature computation, incorporating its own identity information. This process can be formally represented as Eq. (4). Subsequently, a comparison is made with the baseline condition, expressed as $mac_{ij} \stackrel{?}{=} mac_{ij}'$. If the equation holds true, the verification process is deemed successful, warranting further progression. Conversely, if the comparison fails to satisfy the condition, the received user report is deemed invalid and subsequently rejected.

$$mac_{ij}' = h(C_{ij} \| h(T_p \| ID_{ij})) \tag{4}$$

2) Initial data aggregation: Within the purview of this stage, every fog node orchestrates the aggregation of ciphertext originating from all smart meters under its jurisdiction. The resultant aggregated ciphertext is delineated by Formula 5. The generation of the corresponding signature for this aggregated ciphertext is elucidated through Formula 6. Where $y_j = h(T_p \| ID_j)$.

$$C_i = \prod^{n} C_{ij} \bmod N^2 \tag{5}$$

$$mac_j = h(C_j \| y_j) \tag{6}$$

3) Primary node election: In an endeavor to curtail the risk of regional data loss and data compromise stemming from single points of failure, and to mitigate the concentration of processing numerous data streams originating from fog nodes via the measurement data management system, the algorithmic methodology outlined in Algorithm 1 is harnessed for primary node selection. Each of the participating fog nodes is equipped with the prospect of ascending to the role of a master node.

4) Subsequent data aggregation: Following the receipt of the aggregation report denoted as (C_j, mac_j, T_p) dispatched by the respective fog node, the master node undertakes a sequence of actions. Initially, the master node initiates the verification process of the aggregation report. Subsequently, it computes the digital signature, as depicted in Formula 7. In contrast to $mac_{ij} \stackrel{?}{=} mac_{ij}'$, if the equation is valid and T_p is within the validity period, then the verification passes. Then, the master node will perform a secondary aggregation of the aggregation report, and the second-level aggregation ciphertext is C_{AS}, as shown in Formula 8.

$$mac_{ij}' = h(C_{ij} \| h(T_p \| ID_{leader})) \tag{7}$$

$$C_{AS} = \prod_{j=1}^{m} C_j \bmod N^2 \tag{8}$$

5) Creation of new block: The master node is tasked with the creation of transaction $T_x = \left(C_{AS}, Leader_{fog_j}, T_p\right)$. In the network model outlined within this section, all fog nodes bear the responsibility of upholding and overseeing the fog blockchain. However, the role of generating blocks is solely entrusted to one and only one master node. The architectural blueprint of this process is succinctly represented in Fig. 3. Upon the master node's encapsulation of the transaction within a block, it disseminates the block across the network, transmitting it to all nodes. Upon receipt of acknowledgments from over a third of the nodes, the block is seamlessly integrated into the longest blockchain, ensuring its enduring persistence and security.

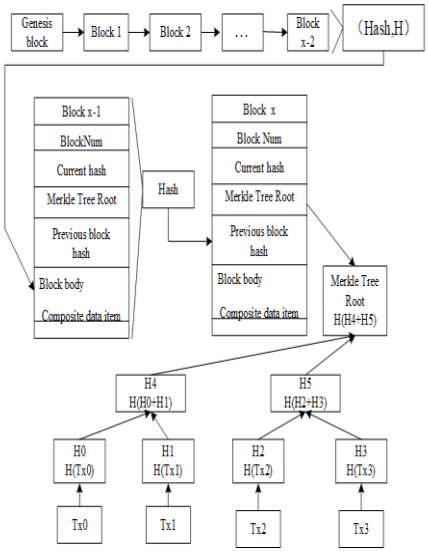

Fig. 3. Block structure and connections.

D. Data Reading and Analysis

Algorithm2 Extraction of Regional Data

Input： M and R

Output： $\left(D_1, D_2, ..., D_n\right)$

1 Set $x_0 = M/R$; $a_1 = R^1, a_2 = R^2, ..., a_m = R^m$; $x_0 = D_1 + R^1 D_2 + ... + R^{m-1} D_m$;

2 For $j = m$; $j > 1$; $j--$ do;

3 $D_j = x_{j-1} \bmod R$;

4 $x_j = x_{j-1}/R$;

5 End for;

6 return $\left(D_1, D_2, ..., D_n\right)$.

The measurement data management system possesses the capability to retrieve information from the fog chain at η minute intervals. Commencing this process, the system initiates the decryption of the second-level aggregate ciphertext through the utilization of the Paillier homomorphic decryption algorithm. To streamline this decryption process for the aggregated ciphertext, certain definitions are introduced:

$$M = a_1 \sum_{i=1}^{n} d_{i1} + a_2 \sum_{i=1}^{n} d_{i2} + ... + a_n \sum_{i=1}^{n} d_{in} \tag{9}$$

$$R = \prod_{j=1}^{n} r_j \tag{10}$$

The ciphertext can then be converted to the form of Formula 11.

$$C = g^M \cdot R^N \bmod N^2 \tag{11}$$

The final aggregated ciphertext still follows the Paillier encryption algorithm, so the measurement data management system can perform Paillier decryption using the private keys λ and $L(\mu)$ to obtain the aggregated plaintext M:

$$M = D(C) = L\left(C^{\lambda \bmod N^2}\right) \cdot \mu \bmod N \tag{12}$$

The ultimate goal of the measurement data management system is to obtain the fine-grained power consumption of each region. In order to achieve this goal, the region data can be obtained through algorithm 2 and $D_1, D_2, ..., D_n$ can be extracted from M:

$$D_j = \sum_{i=1}^{j} d_{ij} \tag{13}$$

VI. Simulation Experiment and Performance Analysis

A. Performance Test

This paper uses Hyperledger Caliper to test the performance of the deployed blockchain network, mainly measuring transaction throughput, transaction latency, and docker container volume. The host hardware parameters used were as follows: The experiment was carried out on an Apple M1 CPU@4*3.2GHz+4*2.064GHz computer.

This paper builds a blockchain network based on Hyperledger fabric v2.4, and the test network consists of two organizations, each with two nodes. A workload named AMI DATA is used to simulate the reading and writing of AMI data, and the performance of the blockchain is evaluated by controlling the number of transactions sent. The test results are shown in Table I.

TABLE I. Performance of Parameters of AMI Data Transaction

AMI data	Maximum delay (s)	Handling capacity (TPS)
1000	1.12	75.4
2000	1.20	79.4
3000	2.05	80.5
4000	1.08	81.6
5000	1.08	98.2

According to the experimental results, with the increase of the number of transactions, the maximum delay basically does not change and remains at about 1S, and the whole system is in

a fairly stable state. The blockchain system can process all the AMI data in a very short time. It can meet the time requirements of smart grid data processing. The throughput changes from 75.4TPS to 98.2TPS, which is a small change and uses less system resources. It can meet the needs of daily power grid equipment. As shown in Fig. 4, the throughput changes with the number of AMI devices. It can be seen that with the increase of AMI devices, the throughput gradually becomes stable. The growth rate is smaller. The changes in throughput are shown in Fig. 4.

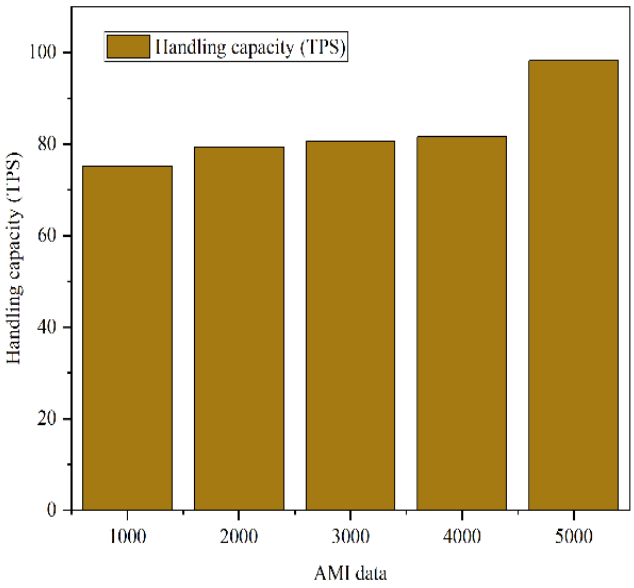

Fig. 4. Changes in throughput.

B. Calculation of Cost

Within this section, a comprehensive analysis of the overall system's computational cost is undertaken. In this analysis, it is posited that the number of fog nodes traverses a spectrum ranging from 5 to 50, which is $m \in [5,50]$. Each of these fog nodes governs jurisdiction over 20 smart meters, which is $n = 20$. For the sake of facilitating a comprehensive performance evaluation of the proposed scheme, a comparative analysis was conducted involving other proposed schemes such as EBDA [20] and LPDA-EC [21]. This comparative assessment aims to provide a robust understanding of the proposed scheme's relative merits and performance attributes in contrast to these alternative approaches. Define T_{E_1} as the exponential operation time in $Z_{n^2}^*$, T_{E_2} as the exponential operation time in G , T_M as the multiplication operation time, and T_P as the pairing operation time. Within this section, the execution of each aforementioned operation is carried out leveraging a pair-based encryption library (PBE). To ensure comparability across experiments, the operation times stipulated in study [13] are adopted. Table II presents the diverse operations alongside their corresponding execution times. Notably, due to the significantly minute time allocation for hash operations relative to other operations, these hash-related procedures are deemed negligible and thus omitted from the computational cost assessment.

TABLE II. TIME TO RUN THE OPERATION

Notations	Descriptions	Time cost （ms）
T_{E_1}	Exponentiation operation in $Z_{n^2}^*$	1.60
T_{E_2}	Exponentiation operation in G	1.62
T_M	Multiplication operation	0.06
T_P	Pairing operation	17.70

At the user level, computing ciphertext C_{ij} requires $2mn$ multiplication T_M . In the fog layer, each fog node performs the first data aggregation, then the entire fog layer requires a total of mn times multiplication operation, that is, mn times T_M . The master node performs a second data aggregation, requiring the m multiplication operation T_M . When the measurement data management system in the cloud needs to read the contents of the fog chain, the operation of Formula 12 is performed, which requires a power operation T_{E_1} and a multiplication operation T_M . In summary, the total calculation cost of the proposed scheme is $(3mn+m+1)T_M + T_{E_1}$. Table III shows a comparison of calculated costs.

TABLE III. COMPARES THE CALCULATED COSTS

Option	Calculate the cost
LPDAEC	$m\left[2nT_{E_1} + 4nT_M + (3n+1)T_{E_2}\right] + 2mT_P + 2mT_{E_1}$
EBDA	$m\left[nT_{E_1} + 2(n+1)T_M\right] + (m+1)T_P + (m+1)T_M + T_{E_1}$
LDASBC	$(3mn+m+1)T_M + T_{E_1}$

Fig. 5 provides a comparative illustration of the cumulative computational costs associated with the three schemes under examination. Evidently, the proposed scheme presented in this paper markedly exhibits significantly lower computational costs in comparison to the other two schemes. This advantageous performance differential becomes even more pronounced with the escalation in the number of smart meters within the system. By comparing the graphs, it can be concluded that LDAS-BC can significantly reduce the computational cost of AMI data and has strong scalability.

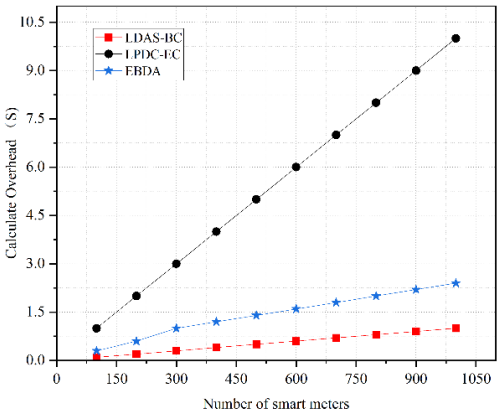

Fig. 5. Calculated cost comparison.

C. Communication Cost

The communication cost within the proposed scheme encompasses two primary components: the communication between the smart meter and the fog node, and the subsequent communication between the fog node and the master node. First, the smart meter generates a user report $\left(C_{ij}, mac_{ij}, T_p, ID_{ij}\right)$ and sends it to the fog node, assuming that $\left|N^2\right|$ is 2048*bit*, $|p|$ is 160*bit*, $|ID|$ and $\left|T_p\right|$ are 160*bit*, then $S_{user\to fog} = \left|C_{ij}\right| + \left|mac_{ij}\right| + \left|T_p\right| = 2368bit$. The fog node then sends $\left(C_{ij}, mac_{ij}, T_p, ID_{ij}\right)$ to the master node, which has $S_{fog\to Leader} = \left|C_{ij}\right| + \left|mac_{ij}\right| + \left|T_p\right| = 2368bit$. As a consequence of the aggregation operation, the aggregate communication overhead remains unaffected by the number of smart grids. This intrinsic property signifies that even as the quantity of smart grids escalates, the communication overhead between the fog node and the master node remains consistent. Table IV tabulates the communication costs associated with the three alternatives.

TABLE IV. COMPARISON OF COMMUNICATION COSTS

Option	Communication Cost
LPDA-EC	$mn\left(\left\|C_{ij}\right\|+160+\left\|T_p\right\|+160\right)+(m-1)\left(\left\|C_j\right\|+\left\|ID_j\right\|+\left\|\sigma_j\right\|+T_p\right)$
EBDA	$mn\left(\left\|C_{ij}\right\|+\left\|h_{ij-s}\right\|+\left\|mac_{ijs}\right\|\right)+(m-1)\left(\left\|C_j\right\|+\left\|ID_j\right\|+\left\|\sigma_j\right\|\right)$
LDAS-BC	$mn * S_{user\to fog}+(m-1) * S_{user\to Leader}$

Fig. 6 presents a comparative analysis of the communication costs inherent in the three schemes. The visual depiction highlights that the proposed scheme boasts a marginal advantage over the other two alternatives in terms of communication costs. This advantage is poised to intensify as the Advanced Metering Infrastructure (AMI) scale expands during subsequent stages. Through the comparison of the figures, it can be concluded that LDAS-BC can significantly reduce the cost of AMI data transmission, and after secondary aggregation, the effect is more obvious.

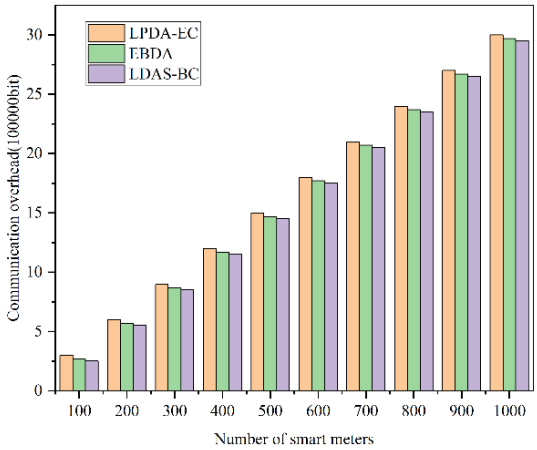

Fig. 6. Calculated cost comparison.

To sum up, compared with existing schemes, LDAS-BC has great advantages in terms of data security, data privacy, data aggregation, as well as computational costs and communication costs. LDAS-BC is a lightweight, low-risk secondary aggregation scheme for blockchain data.

VII. CONCLUSIONS

With the objective of enhancing the data transmission efficiency within the context of advanced measurement systems, as well as curtailing computational and communication overhead, this study introduces the LDAS-BC data aggregation scheme predicated on blockchain technology. This scheme ingeniously amalgamates the Paillier homomorphic encryption algorithm to achieve the aggregation of ciphertext, significantly mitigating the risk of breaching user privacy data. Furthermore, the scheme incorporates the tenets of fog computing and cloud computing, effectively addressing issues tied to computational limitations and restricted storage resources associated with network edge devices. The integration of blockchain technology further reinforces the security and reliability of on-chain data. The proposed LDAS-BC scheme's efficacy is substantiated through theoretical analysis and experimental simulations, which collectively validate its safety and dependability. Comparative experiments additionally underscore the scheme's heightened advantages in comparison to alternative approaches. In the future work, we should consider the selection of consensus algorithm, the mutual authentication among different levels, and the calculation cost and communication cost of these problems need to be calculated. Finally, it is envisaged to extend this scheme to the scenario of Internet of Things data. This will become the focus of the later research work.

REFERENCES

[1] I. Colak, R. Bayindir and S. Sagiroglu, “The Effects of the Smart Grid System on the National Grids,” in 2020 8th International Conference on Smart Grid (icSmartGrid), Paris, France, 2020, doi: 10.1109/icSmartGrid49881.2020.9144891.

[2] K. Sha, N. Alatrash and Z. Wang, "A Secure and Efficient Framework to Read Isolated Smart Grid Devices," IEEE Transactions on Smart Grid, vol. 8, no.6, pp. 2519-2531, 2017, doi: 10.1109/TSG.2016.2526045.

[3] Y.Y. Sun, J.J. Yuan and M.Y. Zhai, “Cloud-Based Data Analysis of User Side in Smart Grid,” in 2016 2nd International Conference on Open and Big Data (OBD), Vienna, Austria, 2016, doi: 10.1109/OBD.2016.13.

[4] P. Zhuang, T. Zamir and H. Liang, "Blockchain for Cybersecurity in Smart Grid: A Comprehensive Survey," IEEE Transactions on Industrial Informatics, vol. 17, no.1, pp. 3-19, 2021, doi: 10.1109/TII.2020.2998479.

[5] H. Tian, Y. Jian and X. Ge, "Blockchain-based AMI framework for data security and privacy protection, " Sustainable Energy, Grids and Networks, vol. 32, pp. 100807, 2022, doi: 10.1016/j.segan.2022.100807.

[6] C. F. Jiang, T. T. Fan, H. H. Gao, W. S. Shi, L. K. Liu et al., "Energy aware edge computing: A survey, " Computer Communications, vol. 151, pp. 556-580, 2020, doi: 10.1016/j.comcom.2020.01.004.

[7] S. L. Chen, H. Wen, J. S. Wu, W. X. Lei, W. J. Hou et al., "Internet of Things Based Smart Grids Supported by Intelligent Edge Computing, " IEEE ACCESS, vol. 7, pp. 74089-74102, 2019, doi: 10.1109/ACCESS.2019.2920488.

[8] L. Chen, R. X. Lu, Z. F. Cao, K. AlHarbi and X. D. Lin, "MuDA: Multifunctional data aggregation in privacy-preserving smart grid communications, " PEER-TO-PEER NETWORKING AND APPLICATIONS, vol. 8, no.5, pp. 777-792, 2015, doi: 10.1007/s12083-014-0292-0.

[9] X. H. Liang, X. Li, R. X. Lu, X. D. Lin and X. M. Shen, "UDP: Usage-Based Dynamic Pricing With Privacy Preservation for Smart Grid, " IEEE TRANSACTIONS ON SMART GRID, vol. 4, no.1, pp. 141-150, 2013, doi: 10.1109/TSG.2012.2228240.

[10] P. Gope and B. Sikdar, "An Efficient Privacy-Friendly Hop-by-Hop Data Aggregation Scheme for Smart Grids" IEEE SYSTEMS JOURNAL, vol. 14, no.1, pp. 343-352, 2020, doi: 10.1109/JSYST.2019.2899986.

[11] P. Singh, M. Masud, M. S. Hossain and A. Kaur, "Blockchain and homomorphic encryption-based privacy-preserving data aggregation model in smart grid, " Computers & Electrical Engineering, vol. 93, pp. 107209, 2021, doi: 10.1016/j.compeleceng.2021.107209.

[12] Z. T. Guan, G. L. Si, X. S. Zhang, L. F. Wu, N. Guizani et al., "Privacy-Preserving and Efficient Aggregation Based on Blockchain for Power Grid Communications in Smart Communities, " IEEE COMMUNICATIONS MAGAZINE, vol. 56, no.7, pp. 82-88, 2018, doi: 10.1109/MCOM.2018.1700401.

[13] S. G. Chen, L. Yang, C. X. Zhao, V. Varadarajan and K. Wang, "Double-Blockchain Assisted Secure and Anonymous Data Aggregation for Fog-Enabled Smart Grid, " Engineering, vol. 8, pp. 159-169, 2022, doi: 10.1016/j.eng.2020.06.018.

[14] F. Loukil, G. G. Chirine, K. Boukadi and A. N. Benharkat, "Privacy-Preserving IoT Data Aggregation Based on Blockchain and Homomorphic Encryption, " Sensors, vol. 21, no. 7, pp. 2452, 2021, doi: 10.3390/s21072452.

[15] C. Regueiro, I. Seco, S. de Diego, O. Lage and L. Etxebarria, "Privacy-enhancing distributed protocol for data aggregation based on blockchain and homomorphic encryption, " Information Processing & Management, vol. 58, no. 6, pp. 102745, 2021, doi: 10.1016/j.ipm.2021.102745.

[16] H. Y. Bao, B. B. Ren, B. B. Li and Q. L. Kong, "BBNP: A Blockchain-Based Novel Paradigm for Fair and Secure Smart Grid Communications, " IEEE INTERNET OF THINGS JOURNAL, vol. 9, no.15, pp. 12984-12996, 2022, doi: 10.1109/JIOT.2021.3107301.

[17] X. H. Zhang, L. You, and G. R. Hu, "An Efficient and Robust Multidimensional Data Aggregation Scheme for Smart Grid Based on Blockchain, " IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, vol. 19 no.4, pp. 3949-3959, 2022, doi: 10.1109/TNSM.2022.3217312.

[18] M. Zhao, Y. Ding, S. J. Tang, H. Liang and H. Y. Wang, "A blockchain-based framework for privacy-preserving and verifiable billing in smart grid, " PEER-TO-PEER NETWORKING AND APPLICATIONS, vol. 16, no.1, pp. 142-155, 2022, doi: 10.1007/s12083-022-01379-4.

[19] R. Y. Yu, A. M. Oguti, D. R. Ochora and S. C. Li, "Towards a privacy-preserving smart contract-based data aggregation and quality-driven incentive mechanism for mobile crowdsensing," Journal of Network and Computer Applications, vol. 207, pp. 103483, 2022, doi: 10.1016/j.jnca.2022.103483.

[20] W. Lu, Z. Ren, J. Xu and S. Chen, "Edge Blockchain Assisted Lightweight Privacy-Preserving Data Aggregation for Smart Grid, " IEEE Transactions on Network and Service Management, vol. 18, no.2, pp. 1246-1259, 2021, doi: 10.1109/TNSM.2020.3048822.

[21] J. Zhang, Y. Zhao, J. Wu and B. Chen, “LPDA-EC: A Lightweight Privacy-Preserving Data Aggregation Scheme for Edge Computing,” in 2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), Chengdu, China, 2018, doi: 10.1109/MASS.2018.00024.

[22] H. B. Fan, Y. N. Liu and Z. X. Zeng, "Decentralized Privacy-Preserving Data Aggregation Scheme for Smart Grid Based on Blockchain," SENSORS, vol. 20, no.18, pp. 5282, 2020, doi: 10.3390/s20185282.