Achieving Efficient and Secure Data Acquisition for Cloud-supported Internet of Things in Smart Grid

Zhitao Guan, Jing Li, Longfei Wu, Yue Zhang, Jun Wu, Xiaojiang Du

Abstract—Cloud-supported Internet of Things (Cloud-IoT) has been broadly deployed in smart grid systems. The IoT front-ends are responsible for data acquisition and status supervision, while the substantial amount of data is stored and managed in the cloud server. Achieving data security and system efficiency in the data acquisition and transmission process are of great significance and challenging, because the power grid-related data is sensitive and in huge amount. In this paper, we present an efficient and secure data acquisition scheme based on CP-ABE (Ciphertext Policy Attribute Based Encryption). Data acquired from the terminals will be partitioned into blocks and encrypted with its corresponding access sub-tree in sequence, thereby the data encryption and data transmission can be processed in parallel. Furthermore, we protect the information about the access tree with threshold secret sharing method, which can preserve the data privacy and integrity from users with the unauthorized sets of attributes. The formal analysis demonstrates that the proposed scheme can fulfill the security requirements of the Cloud-supported IoT in smart grid. The numerical analysis and experimental results indicate that our scheme can effectively reduce the time cost compared with other popular approaches.

Index Terms—Cloud-supported IoT, smart grid, CP-ABE, data acquisition, parallel.

I. INTRODUCTION

With the support of modern information technologies like the Internet of Things (IoT) and cloud computing, smart grid has emerged as the next-generation power supply network, in which the electricity is generated according to the real-time demands of electric equipment or household appliances [1,2]. To make the smart grid more intelligent, a great number of IoT terminals are deployed to gather the status of the power grid timely for the control center. Some sample applications are shown in Fig. 1, such as the power transmission line monitoring, power generation monitoring, substation state monitoring, smart metering, electric energy data acquisition, smart home. For instance, in power transmission line monitoring scenario, using preplaced sensors, the status parameters of the transmission line and power towers can be gathered in real time, so that any fault can be diagnosed and located in a timely manner.

In smart grid, the different kinds of applications mentioned above all generate an enormous amount of data, which needs to be stored and managed efficiently. Cloud-IoT is proposed to address this issue [2,3]. As shown in Fig. 1, with the support of cloud computing, mass data from different IoT terminals can be collected and processed by local front-end servers, then transferred and stored in the cloud servers. The data in cloud can be accessed by various types of data users. The power grid staff can continually monitor the status of power grid. Researchers and government agencies can analyze the data for research or policymaking.

*Jing Li is the corresponding author.
Zhitao Guan, Jing Li and Yue Zhang are with School of Control and Computer Engineering, North China Electric Power University, Beijing, 102206, China (e-mail: guan@ncepu.edu.cn, li_jing@ncepu.edu.cn, minemail@yeah.net).
Longfei Wu and Xiaojiang Du are with the Department of Computer and Information Sciences, Temple University, Philadelphia, USA (e-mail: longfei.wu@temple.edu, dxj@ieee.org).
Jun Wu is with the College of Information Security Engineering, Shanghai Jiao Tong University, Shanghai, China (e-mail: junwuhn@sjtu.edu.cn).

Actually, some works on Wireless Sensor Networks can be used for reference, such as [4-9]. However, there are still several problems and challenges in smart grid data acquisition. First, the efficiency of data acquisition should be considered due to the large amount of data to be encrypted/decrypted and transferred. It’s critical to ensure an acceptable data acquisition time. Second, the protection of data security and privacy must be kept in mind. To deal with these two problems
simultaneously, in this paper, we present an efficient and secure data acquisition scheme based on CP-ABE. The main contributions of our work can be summarized as the following:

✧ We propose a parallel data processing method. Data acquired from the terminals will be partitioned into blocks and encrypted with its corresponding access sub-tree in sequence, thereby the data encryption and data transmission can be processed in parallel. The data decryption process is similar to the process of data encryption.

✧ We introduce the dual secret sharing scheme to protect the access tree information. Only when all of the shares are combined can the secret be recovered. Each of the data blocks holds a share. While the last one share is protected with the other secret sharing scheme. If the user’s attributes satisfy the threshold function of root node, then the last share will be retrieved. In addition, some users with the unauthorized attributes sets will be filtered out. We realize the privacy-preserving, the data integrity check and the attributes check simultaneously.

✧ We give the security analysis and performance evaluation, which prove that the security of our scheme is no weaker than that of the traditional scheme, and that our scheme can reduce the system response time and users’ waiting time notably.

The rest of this paper is organized as follows. Section 2 introduces the related work. The preliminaries are given in section 3. In section 4, the system model and security model are described. In section 5, we present the details of the proposed scheme. Its security analysis and performance evaluation are conducted in section 6 and section 7, respectively. Section 8 concludes the paper.

II. RELATED WORK

Recently, various techniques have been proposed to address the problems of data security and fine-grained access control. In [10], Sahai and Waters proposed the Attribute-Based Encryption (ABE) to realize fine-grained access control on encrypted data. In ABE, the encryption policy is associated with a set of attributes, and the data owner can be offline after data is encrypted. Vipul Goyal et al developed a new cryptosystem for fine-grained sharing of encrypted data in [11] based on Sahai’s work, called Key-Policy Attribute-Based Encryption (KP-ABE). In their scheme, the ciphertext’s encryption policy is associated with a set of attributes, but the attributes that organized into a tree structure (named access tree) are specified by data receivers. In [12], Bethencourt et al proposed the Ciphertext Policy Attribute Based Encryption (CP-ABE). In their work, the data owner constructs the access tree using visitors’ identity information. The user can decrypt the ciphertext only if the attributes in his private key match the access tree.

Owing to the particular advantages of ABE, they are often applied to protect the outsourced data. Yu et al tried to achieve secure, scalable, and fine-grained access control in a cloud environment [13]. Their proposed scheme was based on KP-ABE, and combined with another techniques, proxy re-encryption and lazy re-encryption. It was proved that the proposed scheme can meet the security requirements quite well in cloud. Similarly, [14-16] are other applications of KP-ABE in cloud. Han et al [16] defined a weak anonymity of ABE scheme and proposed a general transformation from ABE to Attribute Based Encryption with Keyword Search.

In [17], the researchers proposed the first KP-ABE scheme enabling truly expressive access structures with constant ciphertext size. In [18], Li et al proposed an expressive decentralizing KP-ABE scheme with constant ciphertext size, which allows the access policy to be expressed as any non-monotone access structure. Similarly, in [19], a fully secure scheme in the standard model is proposed, which is with constant-size ciphertexts and fast decryption simultaneously. There are some other similar researches, such as [20]. The comparison of the above schemes is showed in Table I.

| Constant ciphertext size | Access Structure | KP-ABE/CP-ABE | Security |
|--------------------------|------------------|---------------|----------|
| [17]                     | √                | Non-monotonic | KP-ABE   | CCA-secure |
| [18]                     | √                | Non-monotonic | KP-ABE   | Semantic secure |
| [19]                     | √                | Monotonic     | KP-ABE   | Fully secure in Standard oracle |
| [20]                     | √                | Monotonic     | CP-ABE   | CCA-secure |

To improve system efficiency and protect the user privacy, some researchers study on the multiple authorities. Fu et al [21] firstly presented the Attribute-based Encryption (ABE), secure deletion and secret-sharing schemes and construed a secure multi-authority access control scheme. The scheme proposed in [22] could reduce the reliance on the central authority and protect users’ privacy. The proposed scheme allows each authority to work independently without any collaboration to initialize the system and issue secret keys to users. In [23], Li et al presented a low complexity multi-authority attribute based encryption scheme for mobile cloud computing, which uses masked shared-decryption-keys to ensure the security of decryption outsourcing. They adopted multi-authorities for authorization, to enhance the security assurance.

How to ensure the integrity and correctness of data are challenging issues. In [24], Yadav and Dave presented an
access model based on CP-ABE which could provide a remote integrity check by way of augmenting secure data storage operations. Zhou adopted a similar method [25]. In addition to the access tree division, Zhou also proposed an efficient data management model to balance the communication and storage overhead and reduce the cost of data management operations. Lai et al. [26] gave the formal model of ABE with verifiable outsourced decryption and propose a concrete scheme. Lin et al. [27] proposed a generic construction of verifiable outsourced ABE (VO-ABE), based on an attribute-based key encapsulation mechanism (AB-KEM). It can be considered in both key-policy (KP) and ciphertext-policy (CP) settings. Mao et al. [28] proposed generic constructions of CPA-secure and RCCA-secure ABE systems with verifiable outsourced decryption from CPA-secure ABE with outsourced decryption, respectively.

In [29], M Green et al. proposed a proxy re-encryption in Identity-Based Encryption scheme, in which the ciphertexts can be transformed from one identity to another. In [30], R Canetti et al. proposed a definition of security against chosen ciphertext attacks for PRE schemes. An IBE PRE scheme was proposed in [31] without random oracles. In [32], the first unidirectional proxy re-encryption scheme was proposed with chosen-ciphertext security in the standard model. In [33], SSM Chow et al.’s scheme gains high efficiency and CCA security using the “token-controlled encryption” technique. H Wang et al. realized an identity-based proxy re-encryption scheme that can achieve IND-CCA2 secure [34]. In contrast with requiring users to be online all the time, a time-based proxy re-encryption (Time-PRE) scheme was presented in [35] to revoke a user’s access right automatically after a predetermined period of time. In [36], the proxy re-encryption was proved to be useful as a method of adding access control to the SFS read-only file system. This scheme also achieves a stronger notion of security.

III. PRELIMINARIES

A. Bilinear Maps

Let $G_0$ and $G_1$ be two multiplicative cyclic groups of prime order $p$ and $g$ be the generator of $G_0$. The bilinear map $e$ is $e : G_0 	imes G_0 \rightarrow G_1$ for all $a, b \in \mathbb{Z}_p$:

- Bilinearity: $\forall u, v \in G_1, e(u^a, v^b) = e(u, v)^{ab}$.
- Non-degeneracy: $e(g, g) \neq 1$.
- Symmetric: $e(g^a, g^b) = e(g^b, g^a)$.

B. Discrete Logarithm (DL) Problem:

Let $G$ be a multiplicative cyclic group of prime order $p$ and $g$ be its generator. Given a tuple $<g, g^x>$, where $g \in G$ and $x \in \mathbb{Z}_p$, are chosen as input uniformly at random, the DL problem is to recover $x$.

**Definition 1** The DL assumption holds in $G$ is that no probabilistic polynomial-time (PPT) algorithm $A$ can solve the DL problem with negligible advantage. We define the advantage of $A$ as follows:

$$\Pr[A < g, g^x = x]$$

The probability is over the generator $g$, randomly chosen $x$ and the random bits consumed by $A$.

C. Structure in Ciphertext-policy Attribute Based Encryption (CP-ABE)

**Definition 2** Let $P = \{P_1, P_2, ..., P_n\}$ be a set of participants, let $U = 2^{\{P_1, P_2, ..., P_n\}}$ be the universal set. If $\exists AS \subseteq U \setminus \{\emptyset\}$, then $AS$ can be viewed as an access structure. If $A \in AS, \forall B \in U, A \subset B$, and $B \in AS$, $AS$ is considered a monotonic access structure. Then the sets in $AS$ are defined as authorized sets, while the other sets are regarded as unauthorized sets.

To achieve fine-grained access control, we utilize the Ciphertext Policy Attribute-Based Encryption scheme. For ease of partition, we adopt the Bethencourt’s scheme in [12], in which the access structure is illustrated by an access tree.

Let $T$ be an access tree, and the root node is denoted by $R$. All the leaves represent the attributes, while the interior nodes represent the threshold gates, described by its children and a threshold value, such as AND $(n \text{ of } n)$, OR (1 of $n$), and $n$ of $m \text{ (} m > n \text{)}$.

At the beginning of the encryption, we randomly choose a secret $s$ and conduct a polynomial for each node from top to bottom, while the decryption order is reverse.

Additionally, some functions are necessary to be introduced. We use $\text{parent}(x)$ to get the parent of the node $x$. The function $\text{att}(x)$ is used only if $x$ is a leaf node and returns the attribute associated with the leaf node $x$ in the tree. The function $\text{index}(x)$ returns the number associated with the node $x$, where the index values are uniquely assigned to nodes in the tree.

To retrieve the secret, we define the Lagrange coefficient $\Delta_{xs}$ as follows:

For $i \in \mathbb{Z}_p$, and for $\forall x \in S$,

$$\Delta_{_x,S(x)} = \prod_{j \neq k, j \neq u} \frac{x - j}{i - j}.$$ 

D. $(t, n)$ Threshold Secret Sharing

Secret sharing scheme is used for sharing a secret among a group of parties, each of whom only obtain a piece of the secret (namely a share of the secret). No single party can infer any information about the secret with its own share. The only
way to reconstruct the secret is to combine a certain number of shares.

The most basic secret sharing scheme is \((t, n)\) threshold scheme, which was first proposed by Shamir [37]. In his scheme, a secret divided into \(n\) parts can be recovered only if at least \(t\) parts are collected. This idea has already been used to implement the tree-access structure.

IV. DEFINITIONS

A. Definition of System Model

Our system consists of four entities, namely Data Owner, Cloud Server, Attribute ciphertexts, and Data Requester/Receiver. Both Data Owners (denoted as DO) and Data Requester/Receivers (denoted as DR) are users.

Data Owners (DO) DO decide the access policy and encrypt the data with CP-ABE. The encrypted data will be uploaded to the Cloud Servers. DO are assumed to be honest in the system.

Data Requester/Receivers (DR) DR send the decryption request to Cloud and obtain the ciphertext over the internet. Only when their attributes satisfy the access policies of the ciphertext, can they get access to the plaintexts. Data requester/receivers may collude to access the data that is otherwise not accessible individually.

Cloud Servers (CS) CS are responsible for storing a massive volume of data. They cannot be trusted by DO. Hence, it is necessary for DO to define the access policy to ensure the data confidentiality. CS are assumed not to collude with DR.

Attribute Authority (AA) AA is responsible for registering users, evaluating their attributes and generating their secret key SK accordingly. It runs the Setup algorithm, and issues public key PK and master key MK to each DO. It is considered as fully trusted.

B. Definition of Our System

Definition 3 Our scheme consists of the following algorithms: Setup, Key_Gen, Encryption, and Decryption.

Setup \((1^\lambda) \rightarrow PK, MK\). The setup algorithm takes a security parameter \(\lambda\) as input and outputs the public key and master key.

Key_Gen \((PK, MK, S) \rightarrow SK\). The key generation algorithm takes the public key \(PK\), the master key \(MK\) and a set \(S\) of attributes that belongs to users as inputs. It outputs a corresponding secret key \(SK\).

Encryption \((PK, M, AS) \rightarrow CTB\). The encryption algorithm consists of two subroutines: data partition subroutine Data_Particle and data block encryption subroutine DB_Encryption.

● Data_Particle \((M, AS) \rightarrow DB_1...DB_n\). This subroutine takes the message \(M\) and access structure \(AS\) as inputs. It outputs a number of encrypted data blocks \(DBs\). The number is decided by the threshold function.

● DB_Encryption \((DB_s, AS, PK, MK) \rightarrow CTB_s\). The data block encryption subroutine takes the data blocks \(DBs\), the corresponding sub-access-structure \(AS\), public key \(PK\) and master key \(MK\) as inputs. It outputs the ciphertext blocks \(CTBs\).

Decryption \((DB_s, AS, PK, SK) \rightarrow DB_s\). The data decryption algorithm consists of six subroutines: check of user’s attributes subroutine ATT_Check, check of the ciphertext blocks integrity subroutine CTB_Integrity, the leaf nodes decryption subroutine Decrypt_LeafNode, the interior nodes decryption subroutine Decrypt_InteriorNode, two ciphertext block decryption subroutines CTB_ABE_Dec and CTB_SYM_Dec.

● ATT_Check \((CTB_s, S) \rightarrow R_{s+1}\) or \(\bot\). This subroutine takes as inputs the last ciphertext block \(CTB_s\) and user’s attributes set \(S\). If the check is passed, it returns \(R_{s+1}\). Otherwise, it outputs \(\bot\).

● CTB_Integrity \((CTB.id, R_{s+1}) \rightarrow T\) or \(\bot\). This subroutine takes as inputs all the id of ciphertext blocks and the output of ATT_Check. If the check is passed, it outputs the access tree information \(T\). Otherwise, it outputs \(\bot\).

● Decrypt_LeafNode \((CTB_i, SK, z, T) \rightarrow F_i\) or \(\bot\). This subroutine takes as inputs ciphertext blocks, user’s secret key, leaf nodes information and the access tree \(T\). It outputs \(F_i\) or \(\bot\).

● Decrypt_InteriorNode \((F_i, S_i, T) \rightarrow F_i\) or \(\bot\). This subroutine takes as inputs the outputs of Decrypt_LeafNode, the leaf nodes set \(S_i\) and the access tree \(T\). If \(S_i\) satisfies the current sub-access-policy, it outputs \(F_i\). Otherwise, it outputs \(\bot\).

● CTB_ABE_Dec \((CTB_i, SK, F_i) \rightarrow DB_i\) or \(\bot\). This subroutine takes as inputs \(CTB_i\), user’s \(SK\), and the output of Decrypt_InteriorNode. It outputs \(DB_i\) or \(\bot\).

● CTB_SYM_Dec \((CTB_i, E_i, SK) \rightarrow DB_i\) or \(\bot\). This subroutine takes as inputs the \(CTB_i\), user’s \(SK\) and a parameter \(E_i\). It outputs \(DB_i\) or \(\bot\).

C. Definition of Security Model

Here, we introduce the universal security model of our system, which is defined similar to Bethencourt’s scheme in [12]. In this security model, there is an adversary \(\mathcal{A}\) and a challenger \(\mathcal{C}\). The adversary is allowed to challenge on an encryption to the access structure \(AS^*\) and query for any secret keys \(SKs\) as long as they cannot be applied directly to decrypt
the ciphertext. The challenger is responsible for the ciphertext generation under the $AS^*$ and the secret key generation. Now the security game is described as follows:

**Setup:** The challenger $C$ runs this algorithm. It gives the public parameters $PK$ to the adversary $A$ and keeps $MK$ to itself.

**Phase 1:** $A$ issues queries for repeated private keys corresponding to the sets of attributes $S_1,...,S_q$. ($q$ and $q_1$ are integers randomly chosen by $A$, with $1 < q < q_1$). If any of the sets $S_1,...,S_q$ satisfies the access structure $AS^*$, then aborts. Else, $C$ generates the corresponding secret keys to the sets for $A$. Then he submits a set of attributes and a ciphertext $CT$, and obtains the corresponding $M$ from $C$.

**Challenge:** $A$ submits two equal length messages $M_0$ and $M_1$ to $C$. The challenger $C$ randomly flips a coin $b$, and encrypts $M_b$ under the challenge access structure $AS^*$. Then the generated ciphertext $CT^*$ will be given to $A$.

**Phase 2:** Repeat Phase 1, and the sets are turned from $S_1,...,S_q$ to $S_{q_1},...,S_q$.

**Guess:** The adversary $A$ outputs its guess $b' \in \{0,1\}$ for $b$ and wins the game if $b' = b$.

The advantage of an adversary $A$ in this game is defined as

$$Adv(A) = |Pr[b' = b] - \frac{1}{2}|,$$

where the probability is taken over the random bits used by the challenger and the adversary.

**Definition 4** A CP-ABE scheme is CCA-secure if polynomial time adversaries have at most a negligible advantage in the above game.

**Definition 5** A CP-ABE scheme is CPA-secure if the adversaries cannot make decryption queries in Phase 1.

**V. DESCRIPTION OF OUR SYSTEM**

**A. Overview**

We partition the file and the access tree into blocks, which allows each data block to be encrypted and decrypted independently, in parallel to the transmission and computation of the different data blocks. This reduces the response time of severs and shortens the DR’s waiting time.

As soon as a block is encrypted, it is ready to be transmitted, which allows the next block to be encrypted at the same time. As shown in Fig.2, we use $T_1$ to denote the total time of encryption and transmission of our scheme and $T_2$ denote the time of CP-ABE. The diagram below shows the advantage of our scheme.

**B. Algorithms**

Here we will give the details of the algorithms in our system.

1) **Setup**

This setup algorithm will choose a bilinear group $G_q$ of prime order $p$ with generator $g$. Then, it will choose three random exponents: $\alpha, \beta, \gamma \in _r \mathbb{Z}$. The public key and the master key are published as:

$$PK = \{G_q, g, h = g^\alpha, e(g, g)^\beta\} \quad (1)$$

$$MK = \{\beta, g^{\alpha}, q\} \quad (2)$$

2) **Key_Gen**
This operation is implemented by a specific trusted server independently. It is responsible for registering legal DRs, evaluating the attributes sets and generating the secret keys SKs accordingly. The algorithm is as follows:

**KeyGen**(PK, MK, S) →SK

This algorithm takes as inputs the set of attributes S, public parameters PK and master keys MK. A corresponding secret key is the output. This algorithm first chooses \( r_j \in \mathbb{Z}_p \), \( \forall j \in S \) at random for each attribute. The Hash function \( H_{att} : [0,1]^* \rightarrow G_0 \) is introduced to map any attribute described as a binary string to a random group element. The SK is computed as:

\[
SK = \left\{ D = g^{a_{sr}}, D_j = g^r, D_j' = g^{r_j} \right\}
\] (3)

3) Encryption

This process consists of two main algorithms, **Data_Partition** and **DB_Encryption**. The former one is used for partitioning the original data into several data blocks, the latter one is responsible for encrypting these data blocks by one, which conducts the encryption and transmission in parallel.

**Data_Partition** (M, n) →DB_1...DB_n

DO first constructs an access tree for the given access policy and gets its interior node number n (including the root node). This algorithm takes as inputs M and n, partitioning as follows:

\[
M \rightarrow M_1, M_2, M_3,..., M_{n-1}, M_n
\] (4)

\[
\text{DB}_1 = M_1,
\]

\[
\text{DB}_2 = M_1 \oplus M_2,
\]

\[
\text{DB}_3 = M_2 \oplus M_3,
\]

\[...
\]

\[
\text{DB}_n = M_{n-1} \oplus M_n
\] (5)

**DB_Encryption** (DB_i, AS_i, PK, MK) →CTB_i

If \( i = 1 \), namely that the first block is being encrypted under the first sub-tree. The sub-tree contains only two levels: the parent node and its child nodes. The algorithm performs like following:

It first selects \( s \in \mathbb{Z}_p \) and \( s_i \in \mathbb{Z}_p \) at random and computes:

\[
E_i = g^{\frac{s_i}{s}}
\] (6)

\[
C_i = (DB_i \mid g^{\frac{s}{s_i}}) e(g, g)^{as_i}
\] (7)

\[
C_i = h^r
\] (8)

According to the threshold of the root node, it will construct a polynomial \( q_i(x) \), where \( q_i(0) = s \), which is similar to Bethencourt’s scheme in [12]. If there is any leaf node connecting to the root node, let \( Y_i \) be the set of leaf nodes, it is computed as follows:

\[
\forall y \in Y_i, \hat{C}_{i,y} = g^{q_i(0)}, \hat{C}_{i,y} = H_{att}(\text{att}(y))^{q_i(0)}
\] (9)

Where \( q_i(0) = q_i(\text{index}(y)) \).

In addition, if there is any child node \( y \) being the interior node of the tree, \( q_i(0) = q_i(\text{index}(y')) \). It will be the input for the next sub-tree construction.

Then, a random number \( R_i \) generated by a pseudo-random generator will be considered as the id of CTB_i:

\[
\text{CTB}_i, \text{id} = R_i
\] (10)

The complete form of the first ciphertext block (CTB_i) is as follows:

\[
\text{CTB}_i = (C_1, C_i, \forall y \in Y_i, \hat{C}_{i,y}, \hat{C}_{i,y} \}
\] (11)

If \( i > 1 \), namely that the first data block has been encrypted and transmitted over the internet and the other data blocks are ready for encrypting and transmitting, the encryption process is similar to that of the first data block. It selects \( s_i \in \mathbb{Z}_p \) at random and computes:

\[
C_i = (DB_i \mid g^{\frac{s_i}{s_i}}) e(g, g)^{as_i}
\] (12)

\[
C_i = h^{s_i}
\] (13)

If there is any leaf node belonging to this sub-tree, let \( Y_i \) be the set of leaf nodes, it is computed as follows:

\[
\forall y \in Y_i, \hat{C}_{i,y} = g^{q_i(0)}, \hat{C}_{i,y} = H_{att}(\text{att}(y))^{q_i(0)}
\] (14)

The difference is that \( q_i(0) \) denotes the parent node of the current sub-tree, where

\[
j = \text{parent}(i), q_j(0) = q_j(\text{index}(i)),
\]

\[
i = \text{parent}(y), q_j(0) = q_j(\text{index}(y)).
\]

To guarantee that each of the CTB can be decrypted, the algorithm computes a difference value for this data block:

\[
\Delta C_i = g^{\frac{s_i - q_i(0)}{s_i}} = g^{\Delta s_i}
\] (15)

Similarly, set a random number \( R_i \) to denote the id of CTB_i.

\[
\text{CTB}_i, \text{id} = R_i
\] (16)

The complete form of the ciphertext block CTB_i is as follows:

\[
\text{CTB}_i = (C_1, C_i, \Delta C_i, \forall y \in Y_i, \hat{C}_{i,y}, \hat{C}_{i,y} \}
\] (17)
If \( i = n \), all the relevant computations about \( CTB_i \), \( id, C_a, C_n, \Delta C_a, \forall y \in Y_{a}, \tilde{C}_{a}, \tilde{C}_{n} \)' is similar to the previous operation. In addition, the complete information of the access tree is implicit and protected with \((n+1, n+1)\) secret sharing scheme by randomly generating \( R_i \) and setting
\[
R_{a+1} = T \oplus R_0 \oplus R_1 \oplus \cdots \oplus R_n .
\] (18)

Assume that the threshold function of the root node is \( k \) out of \( t \), then we adopt \((k, t)\) secret sharing scheme to protect \( R_{a+1} \). On one hand, the access policy information is implicit. On the other hand, users with the unauthorized attribute sets will be eliminated.

As we know, the root node \( R \) holds \( t \) branches, and each branch holds several attributes, only when more than \( k \) branches are satisfied can the secret \( s \) be recovered. Thus, all attributes can be divided into \( t \) disjoint sets: \( S_1, S_2, \ldots, S_t \).

The algorithm first constructs a polynomial \( Q(x) \) whose degree is \( k-1 \) and \( Q(0) = R_{a+1} \). Then it will select \( t \) pairs of \((x_i, y_i)_{i \in I} \), and assign them to \( t \) attribute sets. We introduce a one-dimensional array \( A[x] \) and compute as follows:

If \( att_i \in S_1 \), set
\[
A[H_s(att_i)] = H_s(att_i) \oplus (x_i || y_i)
\]
If \( att_i \in S_2 \), set
\[
A[H_s(att_i)] = H_s(att_i) \oplus (x_2 || y_2)
\]
\[
\vdots
\]
If \( att_i \in S_t \), set
\[
A[H_s(att_i)] = H_s(att_i) \oplus (x_t || y_t)
\]
Finally, it outputs:
\[
CTB_a = (C_a, C_n, \Delta C_a, \forall y \in Y_{a}, \tilde{C}_{a}, \tilde{C}_{n} \)' , A
\] (19)

4) Decryption

Once the decryption request is sent, DR’s attributes need to be checked beforehand.

**ATT_Check** \((CTB_a, S) \rightarrow R_{a+1} \) or \( \bot \)

This algorithm takes as inputs the last ciphertext block and user’s attribute set. It can filter out some users with unauthorized sets, which improves the decryption efficiency. It computes as follows:

\[
\forall att_i \in S, A[H_s(att_i)] \oplus H_s(att_i) = x_i || y_i
\] (20)

Finally, it may obtain \( m \) pairs of \((x_i, y_i)_{i \in I} \).

With the polynomial interpolation, if \( m \geq k \), the algorithm can recover the secret and output \( R_{a+1} \); else, it outputs \( \bot \).

**CTB_Integrity** \((CTB.id, R_{a+1}) \rightarrow T \) or \( \bot \)

Next, the integrity of ciphertext will be checked by DR before entering the transmission-decryption mode. This algorithm takes as inputs the \( id \) of all the ciphertext blocks and the output of **ATT_Check**. It computes as follows:

If all ciphertext blocks are available,
\[
CTB_{id} \oplus CTB_{id} \oplus \cdots \oplus CTB_{id} \oplus R_{a+1} = R_0 \oplus R_1 \oplus \cdots \oplus R_n \oplus R_{a+1} = T
\] (21)

Else, this algorithm outputs \( \bot \).

After passing the checks for the attributes and integrity, it will enter the transmission-decryption mode. This algorithm takes as inputs the ciphertext block CTB and SK.

As each sub-tree only has two levels, there will be no recursion algorithm within the calculation. To complete the computation, we define another two algorithms **Decrypt_LeafNode** and **Decrypt_InteriorNode**.

**Decrypt_LeafNode** \((CTB_i, SK, z, T) \rightarrow F_i \) or \( \bot \)

If the node \( z \) is a leaf node, then we let \( j = att(z) \) and perform the computation follows:

If \( j \in S \), then
\[
F_i = e(g^{H_{a+1}(j)}, g^{(0)})
\]
\[
= e(g^{H_{a+1}(j)}, H_{a+1}(att(z))^{(0)})
\]
\[
= e(g, g)^{e_{(0)}}
\]

Else, it returns \( \bot \).

All of the leaf nodes contained in this \( CTB_i \) will be calculated, and the outputs will be stored.

**Decrypt_InteriorNode** \((F_i, S, T) \rightarrow F_i \) or \( \bot \)

If the node \( i \) is the root node of a sub-tree, for all its child nodes \( z \), the algorithm **Decrypt_LeafNode** \((CTB_i, SK, z, T) \) will compute them and store the outputs as \( F_i \). Let \( S \) be the set of child nodes \( z \), if \( \forall z \in S, F_i \bot \) and the function returns \( \bot \); else, we can compute the \( F \) as follows:

\[
F_i = \prod_{z \in S} F_{z, i}^{\Lambda_{z, i} \Lambda_{z, i}^{(0)}}, \text{ where } \left\{ n = \text{index}(z) \atop S_i = \text{index}(z) : z \in S_i \right\}
\]
\[
= \prod_{z \in S} (e(g, g)^{r_{(0)}})^{\Lambda_{z, i} \Lambda_{z, i}^{(0)}}
\]
\[
= \prod_{z \in S} (e(g, g)^{r_{(0)}})^{\Lambda_{z, i} \Lambda_{z, i}^{(0)}}
\]
\[
= e(g, g)^{r_{(0)}}
\] (23)
To guarantee that all of the CTBs can be decrypted even if some of the sub-trees cannot be satisfied, we propose two decryption methods:

**CTB_ABE_Dec** (CTB, SK, F) \(\rightarrow DB_i\) or \(\perp\)

This subroutine takes as input the result of **Decrypt_InteriorNode**. If the result is valid, as long as one of them is retrieved (assume that the parent node \(i\) is retrieved), this algorithm computes as follows:

\[
\frac{C_i}{e(C_i, D)} = \frac{(DB_i \parallel g^{\frac{n}{q}}) e(g, g)^{\alpha q(0)} e(g, g)^{\gamma q(0)}}{e(h^{\gamma q(0)}, g^\frac{\alpha r}{q})} \quad (24)
\]

\[
= (M_{i-1} \oplus M_i) \parallel g^\frac{n}{q}
\]

**CTB_SYM_Dec** (CTB, E, PK, MK) \(\rightarrow DB_i\) or \(\perp\)

Once there is one CTB being decrypted by satisfying its relevant sub-tree, the CTBs that are related to their child nodes can be decrypted legally.

Set \(j = parent(i)\),

\[
E = \frac{E_j}{\Delta C_j} = \frac{g^{s_j} g^{Q_j}}{g^{S_j} g^{Q_j}} = g^{\gamma q(0)}
\]

\[
\frac{C_i e(E, D_i)}{e(C_i, D)} = \frac{(DB_i \parallel g^{\frac{\gamma q}{q}}) e(g, g)^{\alpha q(0)} e(E, g^\gamma)}{e(h^{\gamma q(0)}, g^\frac{\alpha r}{q})} \quad (25)
\]

\[
= (M_{i-1} \oplus M_i) \parallel g^\frac{n}{q}
\]

The message \(M\) can be retrieved as follows:

\[
M_i = M_i,
\]

\[
M_2 = DB_2 \oplus M_1,
\]

\[
M_3 = DB_3 \oplus M_2,
\]

\[\vdots\]

\[
M_n = DB_n \oplus M_{n-1},
\]

\[
M = M_1 \parallel M_2 \parallel M_3 \cdots \parallel M_n
\]

Finally, it outputs \(M\).

VI. SECURITY ANALYSIS

We analyze the security properties of our system considering the security model defined in Section 4.

A. System Security

We parallel the transmission and computation by partitioning the data and access tree into chunks, which reduces the response time of servers and the DR’s waiting time. The ways to encrypt and decrypt will affect the security of the system.

**Theorem 1:** The security of our system is no weaker than that of [12].

**Proof:** We prove this theorem by the following game. Suppose that an adversary \(A\) can attack our scheme with non-negligible advantage.

- **Setup:** The challenger \(C\) runs this algorithm. It gives the public parameters \(PK\) to the adversary \(A\) and keeps \(MK\) to itself.

- **Phase 1:** \(A\) issues queries for the private keys corresponding to the sets of attributes \(S_1,...,S_q\). In addition, he also submits an access structure \(AS^*\). If any of the sets satisfies the access structure \(AS^*\), then aborts. \(C\) generates the corresponding secret keys \(SKs\) to the sets for \(A\).

- **Challenge:** \(A\) submits two equal length messages \(M_0\) and \(M_1\) to \(C\). The challenger \(C\) randomly flips a coin \(b\), and encrypts \(M_b\) under the challenge access structure \(AS^*\). Then the generated ciphertext blocks CTBs will be given to \(A\).

- **Phase 2:** \(A\) issues queries for the private keys as in Phase 1, and the sets are turned from \(S_1,...,S_q\) to \(S_{q+1},...,S_{q'}\).

- **Guess:** The adversary \(A\) outputs its guess \(b' \in \{0,1\}\) for \(b\) and wins the game if \(b' = b\).

Obviously, the game is formulated corresponding to the one in [12]. Thus, if \(A\) can attack our scheme with non-negligible advantage, he can attack the CP-ABE [12] as well.

B. Partition

A complete tree will be partitioned into several sub-trees, and each data block (DB) will be encrypted with the CP-ABE. The difference is that each sub-tree only contains two levels: one root node and its child nodes. In case some ciphertext blocks are not decrypted, we propose two ways for decryption: one is the same as CP-ABE, the other one is to compute the root node value of current sub-tree from its parent block. However, it may lead to the consequence that even if the user’s set doesn’t satisfy the access policy, the user can still perform decryption and obtain some DBs.

**Theorem 2:** Partition doesn’t impact the security of our scheme.

**Proof:** Compared with a complete tree, each level contains partial access policy. There is no other information inserted into these partitioned ones.

As for DBs, all of them are preprocessed for the first one. Without the first one, even all the rests are acquired, none of
them can be retrieved. The only way to access the first DB is to satisfy its access policy, which is the same as that of [12]. Thus, the partition of the message and the tree doesn’t impact the security of the proposed scheme.

C. Privacy-Preserving and Integrity

**Theorem 3:** Our scheme is secure against the adversaries with polynomial time in the length of the access tree information.

**Proof:** To protect the access tree information, we encrypt it with \((n+1, n+1)\) threshold secret sharing scheme, that is, the secret will be shared among \(n+1\) parties. Before encryption, each ciphertext block obtains a randomly generated string as its \(id\). These strings are the \(n\) shares of the secret. The \(n+1\)-th share is implicit, encrypted with another \((k, t)\) threshold secret sharing scheme.

Only with \(n\) valid ciphertext blocks and the recovered \(n+1\)-th share, can the secret be retrieved. However, the strings of attributes are assigned to legal users by the trusted AA. Adversaries who have no knowledge about the attribute strings cannot launch the brute force attack to guess the attribute strings within polynomial time. Thus, they cannot detect the underlying information about the secret without attributes.

**IV. PERFORMANCE EVALUATION**

A. Numerical analysis

In this paper, we partition a large file into several data blocks, which can be encrypted, decrypted, and transmitted concurrently. Therefore, we are able to perform the encryption/decryption and the transmission of different blocks in parallel. Now, we give the efficiency analysis according to the process shown in Fig. 2 and Fig. 3. The time cost in each step is shown in Table II.

| Table II. The Parameters in Performance Evaluation |
|---------------------------------|
| **Symbols** | **Description** |
| \(n\) | Number of blocks M is partitioned into. |
| \(ET_M\) | Encryption time of the whole file |
| \(ET_i\) | Encryption time of data block \(i\). |
| \(TTM\) | Transmission Encryption time of the whole file |
| \(TT_i\) | Transmission time of data block \(i\). |
| \(DT_M\) | Decryption time of the whole file |
| \(DT_i\) | Decryption time of data block \(i\). |

1) **Encryption-Transmission**

The total time of traditional scheme [12] is,

\[
TT_i + \sum ET_i \approx ET_M + TT_n
\]  

We set \(\Delta T\) to denote the difference:

\[
\Delta T \approx ET_M + TT_M - ET_i - TT_n \approx ET_M - ET_i
\]  

For all various data we suppose that there is no delay in the process of block encryption/decryption or transmission. Therefore, there is no time gap between two consecutive blocks in data block encryption, decryption, and transmission. Thus, in either case, \(\Delta T > 0\).

2) **Transmission-Decryption**

The total time of traditional scheme [12] is,

\[
TT_M + DT_M
\]

If \(TT_i > ET_i\), the total time of our scheme is,

\[
\sum TT_i + DT_i \approx TT_M + DT_n
\]

Else, the total time is,

\[
TT_i + \sum DT_i \approx DT_M + TT_i
\]

We set \(\Delta T\) denote the difference:

\[
\Delta T \approx DT_M + TT_M - DT_i - TT_n \approx DT_M - DT_i
\]  

Ideally, \(DT_M > DT_n\), \(TT_M > TT_i\).

Thus, in either case, \(\Delta T > 0\).

B. Experimental results

To evaluate the performance of our system, we implemented a testing environment using the cpabe-toolkit [38].

While most of the computation cost in encryption is due to the access tree, generating polynomials only takes a small portion of the time. We partition the data into blocks so that the encryption and transmission of data blocks can be executed simultaneously. In the comparative test, the different sizes of the text are first encrypted with the same access tree that contains ten levels and a hundred leaf nodes. As different decryption keys can affect the decryption time overhead, we use the same key to conduct the experiment Compared with [12], Fig. 4 (a) and Fig. 4 (b) show the difference of time cost between reference [12] and our scheme during encryption/decryption and transmission.
Fig. 4 (a) Comparison of the encryption and transmission time between CP-ABE and our scheme, when the message size rises.

Encryption and transmission can be concurrently performed in data blocks. Compared to [12], the total time of encryption and transmission has been reduced.

Fig. 4 (b) Comparison of the transmission and decryption time between CPABE and our system, when the number of blocks is different.

Similar to the encryption operation, the transmission and decryption are, also concurrently performed in data blocks. Compared to [12], the total time of decryption and transmission has been reduced.

Next, let the access trees share the same number of interior nodes (i.e., the number of blocks is fixed) but have different numbers of leaf nodes. Fig. 4 (c) shows the difference of time cost between reference [12] and our scheme during encryption. The total time of [12] rises as the number of leaf nodes grows. Although it increases as well in our scheme, the rate is very small and the time is significantly reduced, shown in Fig. 4 (c).

Fig. 4 (c) Comparison of the encryption and transmission time between CP-ABE and our scheme, when the total number of leaves nodes grows.

Finally, with a fixed number of the leaf nodes, the results in Fig. 4 (d) indicate that the total time of our scheme drops slightly as the number of blocks grows.

Fig. 4 (d) Comparison of the encryption and transmission time between CP-ABE and our scheme, when the number of blocks grows.

VI. CONCLUSION

Cloud-IoT techniques are widely deployed in Smart Grid. Huge amount of data is gathered by IoT front-end devices and stored in the back-end cloud servers. However, achieving data security and system efficiency in the data acquisition and transmission process are of great significance and challenging. Existing related schemes cannot deal with this challenging issue well. To tackle with this problem, we propose a secure and efficient data acquisition scheme for Cloud-IoT in smart grid. In the proposed scheme, the large data is partitioned into several blocks, and the blocks are encrypted/decrypted and transmitted in sequence. In addition, we adopt the dual secret sharing scheme, which realizes the privacy-preserving, the data integrity check and the attributes check simultaneously. The analysis shows that the proposed scheme can meet the security requirements of data acquisition in smart grid, and it
also reduces response time overhead significantly compared to other popular schemes. The data of the proposed scheme is not uploaded in real time, it is offline before encryption. The research on data timeliness will be our future work.

ACKNOWLEDGMENT

This work is partially supported by Natural Science Foundation of China under grant 61402171, the Fundamental Research Funds for the Central Universities under grant 2016MS29, as well as the US Air Force Research Lab under grant AF16-AT10 and the Qatar National Research Fund under grant NPRP 8-408-2-172.

REFERENCES

[1] X. Fang, S. Misra, G. Xue, and D. Yang, “Smart grid—The new and improved power grid: A survey”, IEEE communications surveys & tutorials, vol. 14, no. 4, pp. 944-980, Dec. 2012.
[2] J. Singh, T. Pasquier, J. Bacon, and H. Ko, “Twenty Security Considerations for Cloud-Supported Internet of Things”, IEEE Internet of Things Journal, vol. 3, no. 3, pp. 269-284, Dec. 2016.
[3] L. Jiang, L. D. Xu, H. Cai, Z. Jiang, “An IoT-Oriented Data Storage Framework in Cloud Computing Platform”, IEEE Transactions on Industrial Informatics, vol. 10, no. 2, pp. 1443-1451, Feb. 2014.
[4] X. Du, Y. Xiao, M. Guizani, and H. H. Chen, “An effective key management scheme for heterogeneous sensor networks”, Ad Hoc Networks, vol. 5, no. 1, pp. 24-34, Jan. 2007.
[5] X. Du, and H. H. Chen, “Security in wireless sensor networks”, IEEE Wireless Communications, vol. 15, no. 4, pp. 60-66, Aug. 2008.
[6] X. Du, M. Guizani, Y. Xiao, and H. H. Chen, “Defending DoS Attacks on Broadcast Authentication in Wireless Sensor Networks”, in proc. IEEE ICC, 2008, pp. 1653-1657.
[7] X. Hei, X. Du, J. Wu, and F. Hu, “Defending Resource Depletion Attacks on Implantable Medical Devices”, in proc. IEEE GLOBECOM, 2010, pp. 1-5.
[8] X. Hei, and X. Du, “Biometric-based two-level secure access control for Implantable Medical Devices during emergencies”, in proc. IEEE INFOCOM, 2011, pp. 346-350.
[9] W Yu, G Xu, Z Chen, P Moulema, “A Cloud Computing Based Architecture for Cyber Security Situation Awareness”, in proc. International Workshop on Security and Privacy in Cloud Computing (SPCC), 2013, pp. 488-492.
[10] A. Sahai, and B. Waters, “Fuzzy Identity-Based Encryption”, in proc. EUROCRYPT, 2005, pp. 457-473.
[11] V. Goyal, O. Pandey, A. Sahai, and B. Waters, “Attribute-based encryption for fine-grained access control of encrypted data”, in proc. CCS, 2010, pp. 89-98.
[12] J. Bethencourt, A. Sahai, and B. Waters, “Ciphertext-Policy Attribute-Based Encryption”, in proc. IEEE S&P, 2008, pp. 321-334.
[13] S. Yu, C. Wang, K. Ren, and W. Lou, “Achieving secure, scalable, and fine-grained data access control in cloud computing”, in proc. IEEE INFOCOM, 2010, pp. 1-9.
[14] Q. Zheng, S. Xu, and G. Atieniese, “VABKS: Verifiable Attribute-based Keyword Search over Outsourced Encrypted Data”, in proc. IEEE INFOCOM, 2014, pp. 522-530.
[15] X. Liu, J. Ma, J. Xiong, Q. Li, and J. Ma, “Key-Policy Weighted Attribute based Encryption for fine-grained access control.”, in proc. IEEE ICC Workshops, 2014, pp. 694-699.
[16] F. Han, J. Qin, H. Zhao, J. Hu, “A general transformation from KP-ABE to searchable encryption”, Future Generation Computer Systems, vol.30, no. 1, pp. 107-115, Jan. 2014.
[17] N. Attrapadung, B. Libert, E. D. Panafieu, “Expressive Key-Policy Attribute-Based Encryption with Constant-Size Ciphertexts”, in proc. PKC, 2011, pp. 90-108.
[18] Q. Li, H. Xiong, F. Zhang, S. Zeng, “An expressive decentralizing KP-ABE scheme with constant-size ciphertext”, International Journal of Network Security, vol. 15, no. 3, pp. 131-140, May. 2013.
[19] J. Lai, R. H. Deng, Y. Li, and J. Weng, “Fully secure key-policy attribute-based encryption with constant-size ciphertexts and fast decryption”, in proc. CCS, 2014, pp. 239-248.
[20] A. Ge, R. Zhang, C. Chen, C. Ma, “Threshold Ciphertext Policy Attribute-Based Encryption with Constant Size Ciphertexts”, in proc. Australasian Conference on Information Security & Privacy, 2012, pp. 336-349.
[21] Z. Fu, L. Xin, J. Wang, X. Sun, “Data Access Control for Multi-authority Cloud Storage Systems”, in proc. IHI-MSP, 2014, pp. 714-717.
[22] J. Han, W. Susilo, Y. Mu, J. Zhou, “Improving Privacy and Security in Decentralized Ciphertext-Policy Attribute-Based Encryption”, IEEE Transactions on Information Forensics and Security, vol. 10, no. 3, pp. 665-678, Mar. 2015.
[23] F. Li, Y. Rahulamathavan, M. Rajaarajan, C. Phan, “Low Complexity Multi-authority Attribute Based Encryption Scheme for Mobile Cloud computing”, in proc. IEEE SOSE, 2013, pp. 577-579.
[24] H. Yadav, M. Dave, “Secure data storage operations with verifiable outsourced decryption for mobile cloud computing”, in proc. IEEE ICRAIE, 2014, pp. 1-5.
[25] Z. Zhou, D. Huang, “Efficient and secure data storage operations for mobile cloud computing”, in proc. CNSM Workshop on systems virtualization management (svm), 2012, pp. 37-45.
[26] J. Lai, RH. Deng, C. Guan, J. Weng, “Attribute-Based Encryption With Verifiable Outsourced Decryption”, IEEE Transactions on Information Forensics and Security, vol. 8, no. 8, pp. 1343-1354, Aug. 2013.
[27] S. Lin, R. Zhang, H. Ma, M. Wang, “Revisiting Attribute-Based Encryption With Verifiable Outsourced Decryption”, IEEE Transactions on Information Forensics & Security, vol. 10, no. 10, pp. 2119-2130, Oct. 2015.
[28] X. Mao, J. Lai, Q. Mei, K. Chen, “Generic and Efficient Constructions of Attribute-Based Encryption with Verifiable Outsourced Decryption”, IEEE Transactions on Dependable & Secure Computing, IEEE Transactions on Dependable & Secure Computing, vol. 13, no. 5, pp. 533-546, Sept.-Oct. 2016.
[29] M. Green, G. Atieniese, “Identity-Based Proxy Re-encryption”, in proc. ACNS, 2007, pp. 288-306.
[30] C. Ran, S. Hohenberger, “Chosen-ciphertext secure proxy re-encryption”, in proc. CCS, 2007, pp. 185-194.
[31] CK. Chu, WG. Tseng, “Identity-Based Proxy Re-encryption Without Random Oracles”, ISC, 2007, pp. 189-202.
[32] B. Libert, D. Vergnaud, “Unidirectional Chosen-Ciphertext Secure Proxy Re-encryption”, IEEE Transactions on Information Theory, vol. 57, no.3, pp. 1786-1802, Mar. 2011.
[33] SSM. Chow, J. Weng, Y. Yang, RH. Deng, “Efficient Unidirectional Proxy Re-Encryption”, in proc. AFRICACRYPT, 2010, pp. 316-332.
[34] H. Wang, Z. Cao, L. Wang, “Multi-use and unidirectional identity-based proxy re-encryption schemes”, Information Science, vol. 180, no. 20, pp. 4042-4059, Oct. 2010.
[35] Q. Liu, G. Wang, J. Wu, “Time-based proxy re-encryption scheme for secure data sharing in a cloud environment”, Information Science, vol. 258, pp. 355-370, Feb. 2014.
[36] G. Atieniese, K. Fu, M. Green, S. Hohenberger, “Improved proxy re-encryption schemes with applications to secure distributed storage”, ACM (TISSEC), vol 9, no.1, pp. 1-30, Feb. 2006.
[37] A. Shamir, “How to share a secret”, CACM, vol. 22, no. 11, pp. 612-613, Nov.1979.
[38] J. Bethencourt, A. Sahai, B. Waters. “The cpabe toolkit.”, http://acsc.csl.sri.com/cpabe/.