A Privacy Protection Scheme for IoT Big Data Based on Time and Frequency Limitation

Lei Zhang,1,2,3 Yu Huo,1,3 Qiang Ge,2,3 Yuxiang Ma,1,3 Qiqi Liu,3 and Wenlei Ouyang3

1Henan Key Laboratory of Big Data Analysis and Processing, Henan University, Kaifeng 475004, China
2Institute of Data and Knowledge Engineering, Henan University, Kaifeng 475004, China
3School of Computer and Information Engineering, Henan University, Kaifeng 475004, China

Correspondence should be addressed to Yuxiang Ma; y.x.ma@hotmail.com

Received 3 March 2021; Revised 2 April 2021; Accepted 16 June 2021; Published 14 July 2021

Copyright © 2021 Lei Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Various applications of the Internet of Things assisted by deep learning such as autonomous driving and smart furniture have gradually penetrated people’s social life. These applications not only provide people with great convenience but also promote the progress and development of society. However, how to ensure that the important personal privacy information in the big data of the Internet of Things will not be leaked when it is stored and shared on the cloud is a challenging issue. The main challenges include (1) the changes in access rights caused by the flow of manufacturers or company personnel while sharing and (2) the lack of limitation on time and frequency. We propose a data privacy protection scheme based on time and decryption frequency limitation that can be applied in the Internet of Things. Legitimate users can obtain the original data, while users without a homomorphic encryption key can perform operation training on the homomorphic ciphertext. On the one hand, this scheme does not affect the training of the neural network model, on the other hand, it improves the confidentiality of data. Besides that, this scheme introduces a secure two-party agreement to improve security while generating keys. While revoking, each attribute is specified for the validity period in advance. Once the validity period expires, the attribute will be revoked. By using storage lists and setting tokens to limit the number of user accesses, it effectively solves the problem of data leakage that may be caused by multiple accesses in a long time. The theoretical analysis demonstrates that the proposed scheme can not only ensure safety but also improve efficiency.

1. Introduction

The development of emerging computing technologies (e.g., cloud computing) have brought opportunity for various industries, such as hyperspectral remote sensing image algorithms [1, 2], classification algorithms [3], matrix operations under linear systems [4, 5], and data generated by Internet of Things (IoT) devices. If the data in a solution is stored in the cloud or the calculation is outsourced to the cloud, the local storage and calculation pressure will be greatly reduced. Among them, for IoT big data, because IoT devices generate huge amounts of data, the structure of the traditional machine learning model is relatively simple, which can no longer meet the new needs of IoT applications. Thus, deep learning technology has been widely used in IoT applications [6], e.g., smart home [7], smart city [8, 9], and autonomous driving [10].

In the scenario of applying deep learning technology to big data in the IoT, in order to train a neural network, large amounts of data need to be obtained from the IoT devices. For example, crowdsensing systems collect data that comes from sensors embedded on personally owned mobile devices [11]. These data may contain sensitive information of some users. However, IoT networks are becoming more vulnerable to various web attacks [12]. Obviously, once they “share” these IoT data with the same field, they are likely to lose control of this data. If these data containing private information are leaked, and there is a lack of effective protection mechanism...
in the process of IoT search [13], it may cause irreversible harm to the people whose information is leaked. For example, in the field of healthcare, human physiological data collected by wearable IoT devices are put into deep learning models, which can predict the physical condition of patients [14–17]. Once these data are leaked, it will not only cause a patient’s economic loss but also endanger life [18]. In the field of autonomous driving, the prediction system of deep learning may be maliciously interfered. Once location privacy data is obtained maliciously, it may cause traffic safety problems and bring troubles to society [19]. It can be seen that how to protect users’ private data still faces severe challenges for projects that use deep learning to assist IoT applications, and it is a problem that must be solved.

At present, many solutions have been proposed to solve the big data privacy protection problem in machine learning [20] or deep learning. Generally, these schemes are divided into three categories: federated learning [21, 22], encryption-based technologies [23–26], and differential privacy technologies [27, 28], as shown in Figure 1. Figure 1 shows the working principles of three different types of privacy protection. Among them, encryption-based technologies mainly use direct encryption of data, such as using homomorphic encryption algorithms or setting access control on data uploaded to cloud servers. However, in actual situations, data owners not only want to share training data with others but also want to guarantee data security. Although homomorphic encryption solution realizes the encryption of data, it cannot meet the needs of multiuser data sharing when sharing data in the same field, and it cannot achieve one-to-many fine-grained communication. In attribute-based encryption, only users who meet the access strategy set by the owner can obtain the data, which can achieve more flexible access control. Therefore, to handle the problem of the incompatibility of secure storage and fine-grained sharing of IoT big data in deep learning, an attribute-based encryption solution can be introduced. Among them, the encryption of the ciphertext strategy is more suitable to be used in this scenario than the key-based encryption due to the characteristics of the ciphertext contact access strategy and key contact access structure.

In the actual data sharing scenario, due to the numerous attributes of the visitor, there are many departments in the enterprise engaged in the IoT, so the attribute fluidity is relatively large. Access users obtain the key through their own identity attribute information. If the attribute used to represent the identity does not have a valid period, it means that even if an employee resigns or a department merges, it will not affect the access rights of the resigned employee or the original department staff, and these employees can still obtain data through their own identity attributes. If a resigned employee sells IoT big data in exchange for economic benefits, it will not only endanger the interests of the company but also harm people’s personal safety. This shows that it is necessary to set the validity period for each user attribute. The attribute will be cancelled when it expires. Moreover, many current solutions allow users to access unlimited times within the set time. To prevent the number of visits from being abused, it is necessary to limit the number of visits within the set time. To prevent the number of visits from being abused, it is necessary to limit the number of visits within the set time. By limiting the user’s access period and access frequency, to a certain extent, it is possible to reduce the occurrence of data leakage caused by the sale of data information by employees or outsiders using decryption attributes to access big data of the Internet of Things.

We consider the data privacy problems of big data generated in the field of IoT for mobile computing and use attribute revocation idea [29, 30], then propose an IoT big data
privacy protection scheme based on time and the number of decryption restrictions. This scheme combines homomorphic encryption and attribute-based encryption. In summary, the main contributions of this paper are as follows:

(1) We propose a scheme that limits attribute usage time and user decryption frequency. By setting the attribute version number for each attribute as a mark, it is compared with the local time to determine whether the time has expired and realize the revocation. Besides, it limits the number of user accesses by establishing a user decryption frequency table and setting access tokens.

(2) We combine homomorphic encryption with ciphertext-based attribute-based encryption technology, which makes this solution more effective in improving data confidentiality without affecting neural network model training.

(3) We analyse the security of the scheme in a real deployment.

The remainder of the paper is organized as follows. After introducing the related work in Section 2, we provide related technologies used in this paper in Section 3. Section 4 describes the design of our scheme. We analyse security and effectiveness of our scheme in Section 5. Finally, Section 6 concludes this study.

2. Related Work

Although deep learning has brought great convenience to human life, its application is inseparable from data. If some IoT data involves the user’s private information, once it is leaked, it will cause property and life safety issues. More and more solutions [31–34] are proposed to solve data security issues, which are implemented by not directly processing data. In addition, people can also protect their privacy by processing data. Lv et al. [35] proposed a secure transaction framework based on the blockchain, which uses the encryption mechanism of the blockchain to ensure information security, but it does not achieve fine-grained access control. Lindell et al. [36] proposed that two parties can process data sets collaboratively without revealing their privacy. Agrawal et al. [37] proposed a scheme that implements the function of outsourcing data to others for data mining tasks. This scheme is confirmed that it does not reveal the data owner’s private information during the outsourcing process. Homomorphic encryption technology is considered to be the most effective and most direct means of protecting user privacy [38]. It can directly perform operations, and the results can be consistent with the results of plaintext operations. In 2007, Orlandi et al. [39] introduced homomorphic encryption technology and multiparty secure computing technology to feed the encrypted data into the neural network model for training, which not only ensured the consistency of the plaintext and ciphertext calculation results but also considered security. In [40], the authors proposed a neural network model that uses encrypted data for training. At the same time, in this scheme, it is also proved that cloud services can be used to put encrypted data into the neural network for prediction operations, and the results are returned from the cloud in the form of ciphertext. In [41], the authors improved the scheme [40] and proved that encrypted data can also train neural networks.

In addition to directly encrypting big data, there are also many solutions for setting access control to the data protection layer. In [42], the author created the first CP-ABE solutions, the access policy and ciphertext are sent to the receiver together. Due to the existence of user or attribute revocation problems, research on revocation of ABE has always received extensive attention. Shi et al. [43] proposed a scheme based on a hierarchical cryptosystem. Once the attributes are revoked, the public key, private key, and ciphertext of the scheme need to be updated, so the revoking efficiency of this scheme is not high. In [44, 45], the authors pointed out that the private key can be divided into two parts. If the attribute is revoked, the two keys need to be updated, and it is necessary to reencrypt the ciphertext and header files, so the cost of revocation is relatively large. In [46], the authors proposed a user revocation scheme based on a time limit, but it did not achieve fine-grained attribute revocation. In [47], the authors proposed a scheme for using smart contracts to revoke attributes. In addition to these revocation schemes, the purpose of revocation can also be realized by limiting the number of user visits. In [48], the authors proposed a scheme that decryption frequency can be limited. But the function of this scheme is a bit single. While sharing IoT big data that can be used for neural network training, users can adopt a scheme that combines homomorphic encryption and CP-ABE. The solution proposed in [49] has proved that combining the two technologies in such scenarios can not only reduce the risk of data leakage but also reduce the number of key communications. However, in the field of deep learning-assisted IoT applications, there are very few solutions that can combine these technologies to limit user access time and specify the number of user accesses.

3. Preliminaries

3.1. Bilinear Maps. Suppose there is a large prime number p and two cyclic groups $G_1$ and $G_2$, their orders are both p, and g is a generator of $G_1$. Then, there is a mapping $e : G_1 \times G_1 \rightarrow G_2$ from $G_1$ to $G_2$, and it has the following properties [50]:

1. Bilinearity: $e(g^a, g^b) = e(g^a, g^b) = e(g, g)^{ab}$ for $\forall a, b \in \mathbb{Z}_p^*$ and $\forall u, v \in G_1$

2. Nondegeneracy: there exists $x, y \in G_1$, such that $e(x, y) \neq 1$, where 1 is the identity element of group $G_2$

3. Computability: for $\forall u, v \in G_1, e(u, v)$ can be calculated by an effective algorithm.

Then, we call the above mapping e a bilinear mapping. In general, the cyclic group $G_1$ is an additive cyclic group, and the cyclic group $G_2$ is a multiplicative cyclic group.
3.2. Diffie-Hellman Problem. For the additive cyclic group $G$ in the above bilinear map $e$, there are the following difficult problems in cryptography and discrete mathematics, various cryptosystems based on bilinear mapping are built on the basis of these difficult problems.

**Definition 1** (discrete logarithm problem (DL)). If there are any two elements $g$ and $Y$, $g \in G$, $Y \in G$, and satisfy $Y = g^k$, where $k \in Z_p^*$, it is difficult to calculate the value of $k$.

**Definition 2** (computational Diffie-Hellman problem (CDH)). Given that a triplet is $(g, g^a, g^b)$, where $g$ is a generator of group $G$, $a, b \in Z_p^*$, it is difficult to calculate the value of $g^{ab}$.

**Definition 3** (decisional Diffie-Hellman problem (DDH)). If there is a four-tuple $(g, g^a, g^b, g^c)$, where $g$ is a generator, $a, b, c \in Z_p^*$, it is difficult to determine whether $c = ab \mod p$ is true.

Because the above three types of problems are based on group $G$, they are all regarded as group $G$ problems.

3.3. DBDH Assumption. Given that a five-tuple is $(g, g^a, g^b, g^c, X)$, where $g$ is a generator of group $G$, $a, b, c \in Z_p^*$, $X \in G$ is the structure of $G$, it is difficult to determine whether $X = e(g, g)^{abc}$ is true.

3.4. Access Structure. The structure is a set of judgment conditions, usually expressed as $\Gamma$, which contains several attribute elements in the attribute set $A$ and threshold logic operators (such as OR and AND). If there is an attribute set that satisfies the judgment condition, this attribute set is called an authorized set, otherwise, we called it an unauthorized set. Let $P = \{P_1, P_2, \ldots, P_n\}$ be the entity set of $n$ participants.

4. The Proposed System

4.1. System Solution. In our proposed solution, there exist six types of entities: IoT device, cloud server, data user, attribute authorization centre, key generation centre, and time server. The scheme model is shown in Figure 2.

From Figure 2, we can know that the data owner can encrypt all kinds of data from IoT devices and upload the data to CSP. The access user makes an access request to the cloud server. Legitimate users can download document set
from the cloud server and decrypt it. CSP and KGC jointly
generates keys for users through continuous interaction. The
time server is responsible for detecting whether the time sent
to it by other entities has expired or has been forged or tam-
pered with.

4.2. System Algorithms. We let group G be a bilinear group,
let g be a generator in group G. Let e : G × G → G be a bilinear
mapping. We choose three hash functions in this
scheme: \( H : \{0, 1\}^* → G \), so that each attribute can be
mapped to the group, \( H_1 : G → Z_p^* \), and \( H_2 : \{0, 1\}^* → \{0, 1\}^* \). In addition, for any \( i \in Z_p^* \), an attribute set \( A \),
the Lagrange coefficient is defined as \( \Delta_{iA}(x) = \prod_{j \in A, j \neq i}(x - j)/(i - j) \).

(1) Setup(\( \lambda \)) → (PK\(_{\text{KGC}}, MK\(_{\text{KGC}}\)), (PK\(_{\text{CSP}}, MK\(_{\text{CSP}}\)), (PK\(_{\text{sign}}, MK\(_{\text{sign}}\)). First, the security parameter \( \lambda \) is used
to generate three pairs of public and private keys,
which are the key generation centre’s key pair
PK\(_{\text{KGC}}, MK\(_{\text{KGC}}\) the cloud server’s key pair PK\(_{\text{CSP}}, MK\(_{\text{CSP}}\), and the public and private key pair for digital
signature PK\(_{\text{sign}}, MK\(_{\text{sign}}\). KGC randomly selects \( \beta \in G \)
\( Z_p^* \) and sets \( h = g^\beta \), so \( (PK_{\text{KGC}} = h, MK_{\text{KGC}} = \beta) \). At the same
time, KGC also selects a random number \( y \in Z_p^* \), so
that the public and private key pair used for
digital signature is \( (PK_{\text{sign}} = g^y, MK_{\text{sign}} = y) \). CSP randomly
selects \( \alpha \in Z_p^* \), it sets \( (PK_{\text{CSP}} = e(g, g)^\alpha, MK_{\text{CSP}} = g^\alpha) \). Second, CSP allocates initialization
information other than public and private keys for
users accessing IoT data, including setting the unique
identity of the \( i \)th user as \( u_i \), where \( u_i \in Z_p^* \). A list \( L \) is stored in the
cloud server, which contains the user’s unique
mark \( u_i \), the number of user visits \( \sigma \), and
the state-related mark \( K_e \). Third, KGC selects a random
secret value \( r \in Z_p^* \) for the user, and AAC selects
a mark \( v_i \in Z_p^* \) for each attribute. Therefore, the system
public key is \( PK = \{G, g, h, f = g^{1/y}, e(g, g)^\alpha\} \), and
the master key is \( MK = (\alpha, \beta) \). The initial value of \( \sigma \)
is set to 0.

(2) KeyGen(PK, MK, MK\(_{\text{sign}}\), \( \mathcal{A}, U_i, st \)) → (SK\(_{\text{u,i}}\)). In
this part, the digital signature private key MK\(_{\text{sign}}\),
the user’s attribute set \( \mathcal{A} \), and the attribute version
key \( U_i \), and outputs the user’s decryption key. The
following four parts are included:
(a) Generate attribute version key. This part is exec-
uted by AAC. AAC randomly selects any value
\( t_i \in Z_p^* \) for each attribute, and \( t_i \) is used as a
parameter for subsequent use, so the attribute
version key is set to \( U_i = v_i t_i \), and the attribute
version key is generated and sent to CSP.
(b) Generate partial user keys. This part is formed by
the simultaneous operation of KGC and CSP via
introducing a secure two-party computing proto-
col. First, KGC takes the parameters \( (r_j, \beta) \) as
input, and CSP takes the parameter \( \alpha \) as input. Through calculation, \( x = (\alpha + r_j)\beta \) is
obtained, and the result is output to CSP. CSP selects a
random number \( \delta \in Z_p^* \), calculates \( A = g^{\delta} = g^{(\alpha r_j + \beta)\delta} \)
and sends the calculation result to KGC. When
KGC receives the result, calculate \( B = A^{1/\delta} =
(\alpha^{\delta} g^{\beta})^{1/\delta} = g^{\alpha^{\delta} \beta} \), and finally, the result
(4) Encryption(PK, \( \Gamma, M, p \)) → (C\(_{\text{T}}^*\)). This algorithm
first inputs the system public key and access policy
tree \( \Gamma \), homomorphic encryption key \( p \), and plaintext
message \( M \). Then, this algorithm outputs encrypted
ciphertext \( C_{\text{T}}^* \). First, the data owner uses the homomor-
phic encryption key \( p \) to encrypt the plaintext \( M \).
The specific operation is as follows: they choose two
random numbers \( q, r \), where \( p \in [2^{q - 1}, 2^q), r \in [2^{p - 1}, 2^p), p/2 > 2r, q \gg p \). The ciphertext of the
document set is calculated by formula \( pq + 2r + M \), \( M \) is
expressed in binary, and the generated ciphertext is
uploaded. Second, the data owner encrypts the
homomorphic key and uploads the key with attribute
access control to the cloud. The data owner regards
the attributes as leaf nodes, the root node of the tree
is \( R \), and the other nodes are threshold logic oper-
ators. The encryption operation performs from the
root node and, from top to bottom, produces a linked
order for each node, which is \( d \), polynomial \( q_\alpha \). If \( n_\alpha \)
is the threshold of nonleaf nodes, then there is a rela-
tion \( d_\alpha = n_\alpha - 1 \). Then, they select a random value \( s \),
\( \epsilon Z_p \) set the polynomial on the root node to \( q_x(0) = s \), and use the homomorphic encryption key \( p \) to encrypt the plaintext, and use the encryption result to calculate \( C = Enc(p) \cdot e(g, g)^{a_x}, \tilde{C} = H' \). Let the polynomial of other nodes be \( q_x(0) = q_x(x)(index(x)) \), where \( \text{index}(x) \) represents the number associated with any node \( x \). The order of nodes is indicated from left to right. In the entire access policy tree, the information carried by each leaf node must be calculated, \( C_x = g^{q_x(0)}, \quad C_x^* = H(\lambda \cdot q_x(0)) \). Then, the final \( CT^* \) is \( CT^* = \{T, \quad C = Enc(p) \cdot e(g, g)^{a_x}, \tilde{C} = H', \forall \lambda \in \mathcal{F} : C_x = g^{q_x(0)}, \quad C_x^* = H(\lambda \cdot q_x(0)) \} \).

(5) TimeCheck(\( SK_u, S, T, K_r, PK_{\text{sign}} \)). In this part of the algorithm, after the time server receives the validity period of the attribute, it first needs to verify it with digital signature technology to check whether it has been forged or tampered with and verify it with the following calculation method:

\[
e^\left(\frac{H_z(T_i)}{\lambda}, PK_{\text{sign}} \right) = e^\left(\frac{H_z(T_i)}{\lambda}, g_{1|H_z(T_i)} \right) \quad (1)
\]

If the verification is successful, it means that the attribute has not been forged or tampered with. The time server compares the validity period \( T_i \) with the present time to determine whether the attribute has exceeded the validity period. If it has not expired, you need to continue to execute step 6. If it expires, the attribute needs to be revoked. On the contrary, if the verification fails, it means that the validity period \( T_i \) has been maliciously modified, then return \( \bot \).

(6) GenToken(\( u_s, st, ctr_{max} \) \( \rightarrow (B, T) \)): after verifying the attribute validity period \( T_i \), it also needs to verify the user’s access times, but the difference is that even if a certain attribute fails, the user still has the possibility of accessing rights, but if the access times exceed the set threshold, then the user does not have the right to access IoT resource data. This algorithm makes the user’s unique identity \( u_s \), the user’s current state \( st \), and the maximum allowed number of decryption \( ctr_{max} \) into a token and sends the token to the cloud server.

(7) PreDecryption(\( SK_u, B, T \) \( \rightarrow \) timeindex). In this part, the cloud server first detects \( e(g^{R_z(s)} \cdot g^{s_i}, K_p) = E \) and \( K_r = e(g, K_p) \) after receiving the token with information. If it meets the verification conditions, CSP will detect the number of decryption \( \sigma + 1 \leq \text{ctr}_{max} \) in the list \( L \), if it is satisfied, let \( \sigma + 1 \), update \( K_r \) at this time and store it in the list \( L \), and then the user and CSP continue to perform step 8. Then, let timeindex = 1, otherwise, timeindex = \( \bot \). If

(8) Decryption(\( SK_u, CT^* \) \( \rightarrow \) (p)). This part of the algorithm is executed by the decryption user and is divided into the following four parts:

(a) When the node \( x \) in the access policy tree belongs to the leaf node in the access policy tree, let \( i = \text{att}(x) \), it means that the attribute corresponding to the node \( x \) computes

\[
\text{DecryptNode}(SK_u, CT^*, x) = e(D_x, C_x) = e(g, g)^{q_x(0)(\tau_i + U_i)}
\]

\[
e = e(g, g)^{q_x(0)}.
\]

If the attribute is not in the user’s attribute set, return \( \bot \).

(b) When \( \lambda \) belongs to a nonleaf node in the structure tree, we let \( S_z \) be the set of child nodes of each node \( z \) of size \( k_z \). When \( F_z \) exists and the user’s current decryption frequency meet the requirements, then compute

\[
F_x = \prod_{z \in S_z} F_z \cdot \Delta(\text{timeindex}) \Lambda.S_x' = \prod_{z \in S_z} \left( e(g, g)^{q_x(0)} \right)^{\Delta(\text{timeindex}) \Lambda.S_x'}
\]

\[
= \prod_{z \in S_z} \left( e(g, g)^{q_x(0)} \right)^{\Delta(\text{timeindex}) \Lambda.S_x'}
\]

\[
= e(g, g)^{q_x(0)}.
\]

If the root node \( R \) in this structure tree is replaced by the node \( x \) in the above formula, it can be computed as \( A = \text{DecryptNode}(SK_u, CT^*, R) = e(g, g)^{\alpha} \).

(c) When the user’s attribute set meets the requirements, decryption is performed:

\[
\text{Dec} \left( \frac{C}{e(C, D) / A} \right) = \text{Dec} \left( \frac{e(g, g)^{a_x} \cdot Enc(p)}{e(H', g^{(s_{e+1})} / p) e(g, g)^{\alpha}} \right)
\]

\[
= \text{Dec} \left( \frac{e(g, g)^{a_x} \cdot Enc(M)}{e(H', g^{(s_{e+1})} / p) e(g, g)^{\alpha}} \right)
\]

(d) After the data visitor obtains the homomorphic key \( p \), users can obtain the document set by using the homomorphic key \( p \).
5. Safety and Efficiency Analysis

5.1. Solution Security Analysis

(a) First, KGC randomly selects a reencryption parameter $\psi$, which is assigned to AAC, CSP, and users whose attributes have been revoked, so that they can update relevant component information in time. Receiving the update information, AAC updates the attribute version keys $U'_i$ of the revoked attributes that it manages, $U'_i = v_i t'_i$.

(b) The next step is to update the user key. CSP obtains the reencryption parameters allocated in the previous step and regenerates the user’s latest version key together with KGC. The updated user key is $SK_n = \{D = g^{(\alpha r_j)\beta}, D'_\lambda = g^{\alpha f_j} \cdot H(\psi \lambda)^{U'_j}, D'_\lambda = g^{U'_1}, st, K_s, K_p\}$.

(c) The third step is to update the ciphertext. In this part, CSP first selects a random cipher value $s' \in Z^*_p$ to ensure forward security and then updates the relevant components of the ciphertext after receiving the reencryption parameters. The updated ciphertext is

$CT^* = \{\lambda, C = e(g, g)^{\alpha (s+1)}, \text{Enc}_k(M), \tilde{C} = H^s, \forall \lambda \in \mathcal{F}: C_i = g^{\phi_i(0)s'}, C'_\lambda = H(\psi \lambda)^{\phi_i(0)s'}, C'_\lambda = H(\lambda)^{-\phi_i(0)s'}(\lambda \neq \lambda'), D'_\lambda = g^{U'_1}, st, K_s, K_p\}$.

5.1.1. Conclusive security

$5.1.2.$ Forward Security.

The next step is to update the user key. CSP obtains the reencryption parameters allocated in the previous step and regenerates the user’s latest version key together with KGC. The updated user key is $SK_n = \{D = g^{(\alpha r_j)\beta}, D'_\lambda = g^{\alpha f_j} \cdot H(\psi \lambda)^{U'_j}, D'_\lambda = g^{U'_1}, st, K_s, K_p\}$.

(b) The next step is to update the user key. CSP obtains the reencryption parameters allocated in the previous step and regenerates the user’s latest version key together with KGC. The updated user key is $SK_n = \{D = g^{(\alpha r_j)\beta}, D'_\lambda = g^{\alpha f_j} \cdot H(\psi \lambda)^{U'_j}, D'_\lambda = g^{U'_1}, st, K_s, K_p\}$.

(c) The third step is to update the ciphertext. In this part, CSP first selects a random cipher value $s' \in Z^*_p$ to ensure forward security and then updates the relevant components of the ciphertext after receiving the reencryption parameters. The updated ciphertext is

$CT^* = \{\lambda, C = e(g, g)^{\alpha (s+1)}, \text{Enc}_k(M), \tilde{C} = H^s, \forall \lambda \in \mathcal{F}: C_i = g^{\phi_i(0)s'}, C'_\lambda = H(\psi \lambda)^{\phi_i(0)s'}, C'_\lambda = H(\lambda)^{-\phi_i(0)s'}(\lambda \neq \lambda'), D'_\lambda = g^{U'_1}, st, K_s, K_p\}$.
Phase 2: A can always ask B for private key-related information, and then, A guesses the ciphertext and needs to give his own guess value ϑ
c

Guess: if ϑ
c
≠ ϑ′, then DBDH is established, the advantage is

\[ p_1 = p \frac{e(g, g)^{ab}}{e(g, g)} = \epsilon + \frac{1}{2}, \text{ if } p \neq p' \text{, the ciphertext cannot be judged, and the advantage is } p_1 = p' \frac{e(g, g)^{ab}}{e(g, g)} = 1/2. \text{ In summary, } p_1 = (g, g', g^b, g^c, e(g, g)^{abc}) = 1 - p_1 = (g, g', g^b, g^c, e(g, g)^{abc}) = 1 \geq \epsilon. \] It shows that this scheme can realize that no adversary can break the scheme with a nonnegligible advantage in polynomial time.

5.2. Theoretical Comparison. Our scheme is compared with other schemes in terms of revocation mechanism, time limit, number of decryption limits, and anticollusion. The comparison results are shown in Table 1.

From Table 1, it can be seen that in [46–49, 56], the revocation schemes proposed by the authors do not fully meet the revocation needs. Although in [55] the authors proposed a scheme that can support user revocation and attribute revocation, in the scenario we mentioned, it is also a requirement that the ciphertext can be operated. This scheme in [49] realizes that users can operate on ciphertext, but it is not suitable for scenarios where attributes need to be revoked. Our scheme realizes two revocation functions, solves the basic system security problem, and achieves the ciphertext operable function. What is more, we also consider two factors: time and frequency of decryption.

Our scheme is compared with other schemes in terms of key generation efficiency, decryption efficiency, and revocation efficiency. c is the exponential calculation cost, and p is the bilinear pair calculation cost. The comparison results are shown in Table 2.

It can be seen that in [55] only the user performs the decryption operation and in [56] only CSP performs the decryption operation, which will cause one-side pressure. Our scheme effectively reduce the amount of user tasks by placing part of the decryption task on the cloud server. Also, in [55], while realizing user revocation, the cost is np. However, in our scheme, if the user is revoked after judgment, the user only needs to be removed from the list L, thus, its computational complexity is better than the schemes [55, 56]. Although the cost of generating the key is relatively high due to the use of a two-party security protocol, the security of the key is guaranteed through this multiparty cooperation method.

6. Conclusions

Since important personal privacy may be leaked while storing and sharing IoT big data on the cloud, we have proposed an IoT big data privacy protection scheme based on time and decryption frequency limitation, the solution realizes the revocation within the time range and the revocation within the range of decryption times. The access control is set by the combination of homomorphic encryption and attribute-based encryption. In our scheme, legitimate users with a homomorphic encryption key can obtain the original data, and users without a homomorphic encryption key can perform operation training on the homomorphic ciphertext. Our scheme does not only affect the training of the neural network but also improves the confidentiality of the data. At the same time, the security of the system is improved by introducing a secure two-party agreement. Through theoretical analysis, we found that our scheme realizes two revocation functions, solves the basic system security problem, and achieves the ciphertext operable function. While realizing user revocation, the computational complexity is preferable to other schemes. Besides, our scheme can effectively reduce the amount of user tasks by placing part of the decryption task on the cloud server. Therefore, our scheme can not only ensure safety but also improve efficiency. In the next step, we plan to combine the advantages of decentralization and anonymity of blockchain to protect big data in the Internet of Things in a distributed storage environment.
Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

There is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China Project (Nos. 61701170 and U1704122), the Key Scientific and Technological Project of Henan Province (Nos. 202102310340 and 202102210352), the Young Elite Scientist Sponsorship Program by Henan Association for Science and Technology (No. 2020HYTP008), the Foundation of University Young Key Teacher of Henan Province (Nos. 2019GGJS040 and 2020GGJS027), and the Key Scientific Research Project of Colleges and Universities in Henan Province (No. 21A110005).

References

[1] W. Huang, Y. Xu, X. Hu, and Z. Wei, “Compressive hyperspectral image reconstruction based on spatial–spectral residual dense network,” IEEE Geoscience and Remote Sensing Letters, vol. 17, no. 5, pp. 884–888, 2019.

[2] W. Huang, Y. Huang, H. Wang, Y. Liu, and H. J. Shim, “Local binary patterns and superpixel-based multiple kernels for hyperspectral image classification,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 4550–4563, 2020.

[3] L. Peng, H. Zhang, H. Hassan, Y. Chen, and B. Yang, “Accelerating data gravitation-based classification using GPU,” Journal of Supercomputing, vol. 75, no. 6, pp. 2930–2949, 2019.

[4] X. Zhang, F. Ding, and E. Yang, “State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors,” International Journal of Adaptive Control and Signal Processing, vol. 33, no. 7, pp. 1157–1173, 2019.

[5] X. Zhang, L. Xu, F. Ding, and T. Hayat, “Combined state and parameter estimation for a bilinear state space system with moving average noise,” Journal of the Franklin Institute, vol. 355, no. 6, pp. 3079–3103, 2018.

[6] H. Ren, H. Li, Y. Dai, K. Yang, and X. Lin, “Querying in Internet of Things with privacy preserving: challenges, solutions and opportunities,” IEEE Network, vol. 32, no. 6, pp. 144–151, 2018.

[7] E. Park, Y. Cho, J. Han, and S. J. Kwon, “Comprehensive approaches to user acceptance of Internet of Things in a smart home environment,” IEEE Internet of Things Journal, vol. 4, no. 6, pp. 2342–2350, 2017.

[8] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, “Internet of things for smart cities,” IEEE Internet of Things Journal, vol. 1, no. 1, pp. 22–32, 2014.

[9] P. Gope, R. Amin, S. K. Hafizul Islam, N. Kumar, and V. K. Bhalla, “Lightweight and privacy-preserving RFID authentication scheme for distributed IoT infrastructure with secure localization services for smart city environment,” Future Generation Computer Systems, vol. 83, pp. 629–637, 2017.

[10] Y. Tian, K. Pei, S. Jana, and B. Ray, “DeepTest: automated testing of deep-neural-network-driven autonomous cars,” in Proceedings of the 40th International Conference on Software Engineering, pp. 303–314, Gothenburg, Sweden, May 2018.

[11] M. Li, Y. Sun, H. Lu, S. Maharjan, and Z. Tian, “Deep reinforcement learning for partially observable data poisoning attack in crowdsensing systems,” IEEE Internet of Things Journal, vol. 7, no. 7, pp. 6266–6278, 2020.

[12] C. Luo, Z. Tan, G. Min, J. Gan, W. Shi, and Z. Tian, “A novel web attack detection system for Internet of Things via ensemble classification,” IEEE Transactions on Industrial Informatics, vol. 17, no. 8, pp. 5810–5818, 2020.

[13] J. Qiu, Z. Tian, C. Du, Q. Zuo, S. Su, and B. Fang, “A survey on access control in the age of internet of things,” IEEE Internet of Things Journal, vol. 7, no. 6, pp. 4682–4696, 2020.

[14] R. Poplin, A. V. Varadarajan, K. Blumer et al., “Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning,” Nature Biomedical Engineering, vol. 2, no. 3, pp. 158–164, 2018.

[15] H. Li, Y. Yang, T. H. Luan, X. Liang, Z. Zhou, and X. S. Shen, “Enabling fine-grained multi-keyword search supporting classified sub-dictionaries over encrypted cloud data,” IEEE Transactions on Dependable and Secure Computing, vol. 13, no. 3, pp. 312–325, 2016.

[16] G. Cheng, C. Yang, X. Yao, L. Guo, and J. Han, “When deep learning meets metric learning: remote sensing image scene classification via learning discriminative CNNs,” IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 5, pp. 2811–2821, 2018.

[17] A. Rachedi and A. Benslimane, “Multi-objective optimization for security and QoS adaptation in wireless sensor networks,” in 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, May 2016.

[18] W. Zhou, Y. Jia, A. Peng, Y. Zhang, and P. Liu, “The effect of IoT new features on security and privacy: new threats, existing solutions, and challenges yet to be solved,” IEEE Internet of Things Journal, vol. 6, no. 2, pp. 1606–1616, 2018.

[19] M. Amoozadeh, A. Raghuramu, C.-n. Chuah et al., “Security vulnerabilities of connected vehicle streams and their impact on cooperative driving,” IEEE Communications Magazine, vol. 53, no. 6, pp. 126–132, 2015.

[20] M. Shafiq, Z. Tian, A. K. Bashir, X. du, and M. Guizani, “CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine learning techniques,” IEEE Internet of Things Journal, vol. 8, no. 5, pp. 3242–3254, 2021.

[21] G. Xu, H. Li, S. Liu, K. Yang, and X. Lin, “VerifyNet: secure and verifiable federated learning,” IEEE Transactions on Information Forensics and Security, vol. 15, pp. 911–926, 2019.

[22] L. Jiang, X. Lou, R. Tan, and J. Zhao, “Differentially private collaborative learning for the IoT edge,” in Proceedings of the 2019 International Conference on Embedded Wireless Systems and Networks (EWSN ’19), Juncture Publishing, USA, 2019.

[23] M. Hao, H. Li, X. Luo, G. Xu, H. Yang, and S. Liu, “Efficient and privacy-enhanced federated learning for industrial artificial intelligence,” IEEE Transactions on Industrial Informatics, vol. 16, no. 10, pp. 6532–6542, 2019.

[24] G. Xu, H. Li, Y. Dai, K. Yang, and X. Lin, “Enabling efficient and geometric range query with access control over encrypted spatial data,” IEEE Transactions on Information Forensics and Security, vol. 14, no. 4, pp. 870–885, 2019.
[25] H. Li, D. Liu, Y. Dai, T. H. Luan, and S. Yu, “Personalized search over encrypted data with efficient and secure updates in mobile clouds,” *IEEE Transactions on Emerging Topics in Computing*, vol. 6, no. 1, pp. 97–109, 2018.

[26] X. Li, S. Liu, F. Wu, S. Kumari, and J. J. P. C. Rodrigues, “Privacy preserving data aggregation scheme for mobile edge computing assisted IoT applications,” *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4755–4763, 2018.

[27] N. Pappas, S. Song, I. Mironov, A. Raghunathan, K. Talwar, and U. Erlingsson, *Scalable Private Learning with PATE*, 2018.

[28] C. Xu, J. Ren, L. She, Y. Zhang, Z. Qin, and K. Ren, “EdgeSanitizer: locally differentially private deep inference at the edge for mobile data analytics,” *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5140–5151, 2019.

[29] J. Hur, “Improving security and efficiency in attribute-based data sharing,” *Transaction on Knowledge And Data Engineering*, vol. 25, no. 10, pp. 2271–2282, 2013.

[30] J. Y. Wang and X. J. Zhou, “An attribute-based encryption scheme for ciphertext policy that supports attribute revocation,” *Computer Engineering*, pp. 1–7, 2020.

[31] Y. Sun, Z. Tian, M. Li, S. Su, X. Du, and M. Guizani, “Honeypot identification in softwarized industrial cyber-physical systems,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5542–5551, 2020.

[32] Y. Pang, L. Peng, Z. Chen, B. Yang, and H. Zhang, “Imbalanced learning based on adaptive weighting and Gaussian function synthesizing with an application on android malware detection,” *Information Sciences*, vol. 484, pp. 95–112, 2019.

[33] R. Ravindranath, S. Ramasamy, R. Somula, K. S. Sahoo, and A. H. Gandomi, “Swarm intelligence based feature selection for intrusion and detection system in cloud infrastructure,” in *2020 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1–6, Glasgow, UK, July 2020.

[34] D. Xu, J. Pan, X. Du, B. Wang, M. Liu, and Q. Kang, “Massive fishing website URL parallel filtering method,” *IEEE Access*, vol. 6, pp. 2378–2388, 2018.

[35] L. Lv, Z. Yang, L. Zhang, Q. Huang, and Z. Tian, “Multi-party transaction framework for drone services based on alliance blockchain in smart cities,” *Journal of Information Security and Applications*, vol. 58, no. 4, p. 102792, 2021.

[36] Y. Lindell and B. Pinkas, “Privacy preserving data mining,” in *Proceedings of the 20th Annual International Cryptology Conference on Advances in Cryptology*, Springer, Berlin, Heidelberg, 2000.

[37] R. Agrawal and R. Srikant, “Privacy-preserving data mining,” in *Proceedings of the 2000 ACM SIGMOD international conference on Management of data - SIGMOD ’00*, Dallas, Texas, USA, 2000.

[38] T. Plantard, W. Susilo, and Z. Zhang, “Fully homomorphic encryption using hidden ideal lattice,” *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 12, pp. 2127–2137, 2013.

[39] C. Orlandi, A. Piva, and M. Barni, “Oblivious neural network computing via homomorphic encryption,” *EURASIP Journal on Information Security*, vol. 2007, no. 1, Article ID 037343, 2007.

[40] N. Dowlin, G. B. Ran, K. Laine, K. Lauter, M. Naehrig, and J. Wernsing, “CrypNets: applying neural networks to encrypted data with high throughput and accuracy,” in *International Conference on Machine Learning*, pp. 201–210, NY, USA, 2016.

[41] E. Hesamifard, H. Takabi, and M. Ghasemi, “Privacy-Preserving Machine Learning in Cloud,” in *Proceedings of the 2017 on cloud computing security workshop*, pp. 39–43, 2017.

[42] M. Chase, “Multi-authority attribute based encryption,” in *Conference on Theory of Cryptography*, Springer-Verlag, 2007.

[43] J. Shi, C. Huang, K. He, and X. Shen, “ACS-HCA: an access control scheme under hierarchical cryptography architecture,” *Chinese Journal of Electronics*, vol. 28, no. 1, pp. 52–61, 2019.

[44] J. Li, W. Yao, J. Han, Y. Zhang, and J. Shen, “User collusion avoidance CP-ABE with efficient attribute revocation for cloud storage,” *IEEE Systems Journal*, vol. 12, no. 2, pp. 1767–1777, 2017.

[45] S. Wang, K. Guo, and Y. Zhang, “Traceable ciphertext-policy attribute-based encryption scheme with attribute level user revocation for cloud storage,” *PloS One*, vol. 13, no. 9, article e0203225, 2018.

[46] G. Dlixat, S. Y. Han, A. Gulmira, and H. Nurramat, “Time-based user revocation CP-ABE scheme,” *Journal of Xinjiang University(Natural Science Edition)*, vol. 36, no. 3, pp. 324–329, 2019.

[47] X. Qin, Y. Huang, Z. Yang, and X. Li, “An access control scheme with fine-grained time constrained attributes based on smart contract and trapdoor,” in *2019 26th International Conference on Telecommunications (ICT)*, Hanoi, Vietnam, April 2019.

[48] J. Ning, Z. Cao, X. Dong, K. Liang, H. Ma, and L. Wei, “Auditable σ-time outsourced attribute-based encryption for access control in cloud computing,” *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 1, pp. 94–105, 2018.

[49] Y. Tan, L. Lu, and J. Y. Wang, “Ciphertext-policy attribute encryption scheme based on homomorphic encryption,” *Computer Engineering and Applications*, vol. 55, no. 19, pp. 115–120, 2019.

[50] J. Li, Y. Wang, Y. Zhang, and J. Han, “Full verifiability for outsourced decryption in attribute based encryption,” *IEEE Transactions on Services Computing*, vol. 13, no. 3, pp. 478–487, 2020.

[51] Y. Tang and D. Y. Xu, “A secure two-party computation problem based on the convolution,” *Journal of Guizhou University(Natural Ences)*, vol. 33, no. 1, pp. 52–57, 2016.

[52] M. Chase and S. S. M. Chow, “Improving privacy and security in multi-authority attribute-based encryption,” in *Proceedings of the 16th ACM conference on Computer and communications security - CCS ’09*, pp. 121–130, Chicago, Illinois, USA, 2009.

[53] S. S. M. Chow, “Removing escrow from identity-based encryption,” in *International Workshop on Public Key Cryptography*, Springer, Berlin, Heidelberg, 2009.

[54] Y. Dodis and A. Yampolskiy, “A verifiable random function with short proofs and keys,” in *Public Key Cryptography - PKC 2005*, Springer, 2005.

[55] X. Li, S. Tang, L. Xu, H. Wang, and J. Chen, “Two-factor data access control with efficient revocation for multi-authority cloud storage systems,” *IEEE Access*, vol. 5, pp. 393–405, 2017.

[56] Z. T. Jiang, J. Huang, S. Hu, and Z. Xu, “Fully-outsourcing CP-ABE scheme with revocation in cloud computing,” *Computer Science*, vol. 46, no. 7, pp. 114–119, 2019.