A SURVEY REPORT ON TECHNIQUES FOR DATA CONFIDENTIALITY IN CLOUD COMPUTING USING HOMOMORPHIC ENCRYPTION

Dhruka Gaidhani  
Dept. of Computer Science  
Fr. Conceicao Rodrigues College of Engineering  
Mumbai, India

Joshua Koyeerath  
Dept. of Computer Science  
Fr. Conceicao Rodrigues College of Engineering  
Mumbai, India

Neel Kudu  
Dept. of Computer Science  
Fr. Conceicao Rodrigues College of Engineering  
Mumbai, India

Prof. Mahendra Mehra  
Dept. of Computer Science  
Fr. Conceicao Rodrigues College of Engineering  
Mumbai, India

Abstract: In order to store information, access content ubiquitously and run processes remotely, without the concern of existing infrastructure, a majority of users have resorted to using Cloud based technology. Cloud computing uses shared pools of configurable resources to provide varying computational capabilities to store and process data. This service is broadly classified into private clouds and third-party operated clouds. Because of the shared nature of this technology, cloud computing poses confidentiality concerns. Important data might be leaked either by the service provider itself, accidentally or intentionally, or by an external attacker who manages to gain unauthorized access. The solution is data encryption. However, encrypting data will make it unusable for computations or processing. Homomorphic Encryption techniques allow encrypted data to be processed as if plain text data to give the same output. Thus the service provider will have access to encrypted data and can perform operations on it as well; however, the original data will be unknown. In this paper, we aim to present a survey on homomorphic encryption techniques from different categories and a comparative analysis of their application to maintain data confidentiality in cloud computing.

Keywords: Cloud Computing, Cloud Security, Confidentiality, Cryptography, Homomorphic Encryption Algorithms, Integrity

I. INTRODUCTION

CLOUD computing has gained increasing popularity because of the wide range of services it provides. Apart from Infrastructure-as-a-service (IaaS), which is providing on-demand physical computer resources, Platform-as-a-Service (PaaS), which is providing appropriate environment to application developers and Software-as-a-Service (SaaS), which is providing the user with application software and databases, Cloud computing also features big data analytics, disaster recovery and cloud based backup. However, incidents like the 2014 Dropbox security breach and the iCloud leak [1] critically call into question the confidentiality of the user data stored in the Cloud; accessible to the service provider.

A solution to this issue is encryption of data prior to the storage. However, while encryption satisfies security constraints it largely reduces the usability of the data. Thus computations cannot be performed and the cloud is reduced to a remote storage. To circumvent this issue, we use Homomorphic Encryption Schemes. These techniques allow us to use encrypted data as an input to various processes. The results so obtained are also encrypted and can only be deciphered by the designated receiver who has the key [2]. Thus, the confidentiality of the original data is maintained.

However, while there are multiple variations of Homomorphic Encryption algorithms; performing arbitrary operations over encrypted data incurs a large penalty in terms of overhead. Time and computational power needs to be redirected for this. Hence, there is a need to identify the appropriate application of the various available Homomorphic Encryption based algorithms and evaluate them based on their feasibility of use.

In section 2 we briefly discuss about the background of Homomorphic encryption algorithms. In section 3 we discuss the functions and properties of homomorphic encryption algorithms. We move on to explain the available algorithms according to the respective categories. In section 4 we provide a comparative analysis of the aforementioned techniques. In section 5 we discuss possible paths research related to this topic might take. Finally, we conclude the paper in section 6.

II. HOMOMORPHIC ENCRYPTION

Homomorphic Encryption is the conversion of plain text data into cipher text that can be analyzed and processed as if it were in its original unencrypted form. Mathematically, we say that an encryption system is homomorphic if:

From Enc(x) and Enc(y), it is possible to calculate Enc(f(x, y)), without using the private key of the sender, where f can be:

\[ f(x, y) = \{+\text{(addition)}, \times (\text{multiplication}), \oplus (\text{XOR})\} \]

This concept was first put forth by Ronald Rivest, Leonard Adleman and Michael Dertuzos in 1978. In this section we will first discuss the extant properties of homomorphic encryption, followed by the general process of homomorphic encryption and some important classifications. The next section will be based on the categorization provided in this
Properties of Homomorphic Encryption

We broadly categorize these into two:

Additive Homomorphic Encryption: Mathematically this can be described as [4]:

\[ \text{Ek}(PT_1 \oplus PT_2) = \text{Ek}(PT_1) \oplus \text{Ek}(PT_2) \]

Additive homomorphic encryption allows the following identities:

i) The product of two cipher texts will decrypt to the sum of their corresponding plaintexts,

\[ D(\text{E}(m_1, r_1) \cdot \text{E}(m_2, r_2) \mod n_2) = m_1 + m_2 \mod n. \]

ii) The product of a cipher text with a plaintext raising o will decrypt to the sum of the corresponding plaintexts,

\[ D(\text{E}(m_1, r_1) \cdot m_2^o \mod n_2) = m_1 + m_2 \mod n. \]

Multiplicative Homomorphic Encryption: Homomorphic encryption is multiplicative, if [5]:

\[ \text{Ek}(PT_1 \otimes PT_2) = \text{Ek}(PT_1) \otimes \text{Ek}(PT_2) \]

Categorization of Homomorphic Encryption

We classify the algorithms based on the above mentioned properties into

- Partially Homomorphic Encryption (PHE): Permits operations on encrypted data like either multiplication or addition, but not both [6].

- Somewhat Homomorphic Encryption (SWHE): Permits more than one operation – multiplication and addition, but the number of operations is limited [6].

- Fully Homomorphic Encryption (FHE): Permits multiple – multiplication and addition operations without a restriction on the number of operations [6].

General process

The general process of the Homomorphic Encryption systems is: [7], [8], [9].

- Key Generation: The client generates a public key (pk) and a private key (sk).
- Encryption: The client encrypts data with encryption key (pk or pk+sk).
  \[ C_i = \text{E}(\text{Sk}(\text{Pk})) \]
- Storage: This encrypted data \( C_i \) and pk, are stored in the cloud database.
- Request: The client requests the server to retrieve the data or to perform operations on encrypted data.
- Processing: The processing server processes the request and performs the operations requested by the client. This as per the required function using Pk.

Response: Cloud provider returns to the client the processed result.
Decryption: The client decrypts the returned result, using sk.

\[ \text{Result} = \text{Dsk} (\text{EvPk}(f, \text{ESk}(a), \text{ESk}(b))) \]

\[ \text{Result} = \text{Dsk} (\text{EvPk}(f, C_{11}, C_{12})) \]

III. HOMOMORPHIC ENCRYPTION TECHNIQUES

Of the various techniques available we shall present at least one of each category. We shall also specify the operation that is permitted in each algorithm. In the next section we compare our findings.

A. RSA – Multiplicative PHE

Rivest, Shamir and Adleman published their public key cryptosystem in 1978 [10]. Although it is a very basic algorithm it is one of the most crucial building blocks of homomorphic encryption which is why it has been included as an example of multiplicative partial homomorphic encryption technique. Other techniques include El Gamal.

1. Key Generation

   Step 1: User creates a public/private key pair by selecting, two large primes at random - p, q

   Step 2: System modulus N is calculated = p.q

   Euler’s totient \( \phi(N) = (p-1)(q-1) \)

   Step 3: selecting at random the encryption key e

   Where, \( 1 < e < \phi(N) \) \( \gcd(e,\phi(N)) = 1 \)

   Step 4: Public key: PK = \{e,N\}

   Private key: SK = \{d,p,q\}

2. Encryption

   Step 1: To be used – public key PK = \{e,N\}

   Step 2: computes: \( C = M^e \mod N \), where \( 0 \leq M \)

   Encrypted text C is obtained.

3. Decryption

   Step 1: To be used - private key SK = \{d,p,q\}

   Step 2: computes: \( M = C^d \mod N \) to get back the original plain text message

Figure 1: RSA Algorithm

Following figure shows the homomorphic property of the RSA.

Suppose there are two cipher texts, CT1 and CT2.

\[ CT_1 = m_1^e \mod n \]

\[ CT_2 = m_2^e \mod n \]

\[ CT_1 \cdot CT_2 = (m_1^e \cdot m_2^e) \mod n \]
Thus, multiplicative property: \((m_1 m_2)^e \mod n\) [9] is displayed.

It is apparent that RSA is a basic algorithm that provides us with limited computation options. This greatly reduces actual practical applications of this algorithm. However, it is a very important algorithm because it acts as a building block and many enhanced algorithms are based on RSA or use it for some part of the implementation. It is also important to note that comparatively RSA is fast and can be feasibly implemented. An alternative to RSA is the El Gamal algorithm [10].

### B. Goldwasser-Micali System (GM) – AdditivePHE

The Goldwasser-Micali(GM) system is a probabilistic (asymmetric) public-key encryption scheme, developed by Shafi Goldwasser and Silvio Micali in 1982. It is an additive Homomorphic Encryption, but it can encrypt just a single bit [3][11]. It provides data confidentiality however it is not efficient in terms of space complexity because in several cases the cipher text generated is many times larger than the input plain text. GM algorithm as shown in the figure

<table>
<thead>
<tr>
<th>1. Key Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: Each user generates a public/private key pair by selecting, two large primes at random - p, q</td>
</tr>
<tr>
<td>Step 2: computing their system modulus N = p.q</td>
</tr>
<tr>
<td>Choose b € B_n</td>
</tr>
<tr>
<td>Step 3: the key pair is obtained as (pk, sk)</td>
</tr>
<tr>
<td>Where pk = (n, b)</td>
</tr>
<tr>
<td>And sk = (p, q)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Encryption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input message M is made up of multiple bits m1,m2…m3</td>
</tr>
<tr>
<td>Public key is pk PT=CT=Zn</td>
</tr>
<tr>
<td>Step 1: For every bit mi, a random value yi is generated</td>
</tr>
<tr>
<td>Step 2: compute: (C_i = y_i^2 \cdot x_i^m \mod N)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Decryption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input message C is made up of multiple bits c1,c2…c3</td>
</tr>
<tr>
<td>Private key is sk</td>
</tr>
<tr>
<td>Step 1: calculate: (e_i = (c_i/p))</td>
</tr>
<tr>
<td>Step 2: if (e_i = 1) then (m_i = 0) else (m_i = 1)</td>
</tr>
<tr>
<td>Output message is M = m1,m2,…m3</td>
</tr>
</tbody>
</table>

**Figure 2: Goldwasser-Micali Algorithm**

Assume we have to encrypt two bit: m1 and m2 using the Goldwasser Micali cryptosystem.

Additive: \(\text{EncGM}(m_1) \times \text{EncGM}(m_2) \equiv (b^{m_1} \cdot r_1^2 \cdot (b^{m_2} \cdot r_2^2)) \mod n\)

\(\equiv b^{m_1+m_2} (r_1r_2)^2 \mod n\)

\(\equiv \text{EncGM}(m_1\oplus m_2, \text{pk}) [3]\)

### C. The Sander-Young-Yung Cryptosystem (SYY) – SWHE

Circuits involving OR and NOT gates in a bidirectional communication have been put forth by Sander, Young and Yung. In these studies one party is assumed to have knowledge of a circuit computing some secret function f, and the other has a secret input x for which it would like to learn this function and its behavior. Their protocol could be used to calculate AND by DeMorgan’s theorem. They also present a separate AND-homomorphic cryptosystem. This system takes a major inspiration from the Goldwasser-Micali cryptosystem. [12]

<table>
<thead>
<tr>
<th>1. Key Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: security parameter : e</td>
</tr>
<tr>
<td>a positive non-zero integer: l</td>
</tr>
<tr>
<td>Step 2: Create an instance of GM Cryptosystems using security parameter e</td>
</tr>
<tr>
<td>Gen(e, l) = (n, p, q, m)</td>
</tr>
<tr>
<td>Public key = (n, m, l)</td>
</tr>
<tr>
<td>Private key = p</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Encryption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input message M is made up of multiple bits m1,m2…m3</td>
</tr>
<tr>
<td>Public key is pk PT=CT=Zn</td>
</tr>
<tr>
<td>Step 1: Enc(pk, m) generates a vector (v \in (Z_2)^l) to return cipher text c</td>
</tr>
<tr>
<td>Step 2: compute: (c = E(pk, v) = (E(pk, v_1), E(pk, v_2),..., E(k, vl)))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Decryption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input message C is made up of multiple bits c1,c2…c3</td>
</tr>
<tr>
<td>Private key is sk</td>
</tr>
<tr>
<td>Step 1: calculate: (v = D(sk, c) = (D(sk, c_1), D(sk, c_2),..., D(sk, cl)))</td>
</tr>
<tr>
<td>Where D is the decryption function from GM cryptosystem</td>
</tr>
<tr>
<td>Step 2: M = Decode(v).</td>
</tr>
<tr>
<td>Output message is M = m1,m2,…m3</td>
</tr>
</tbody>
</table>

**Figure 3: SYY Algorithm**

The homomorphic operation of the Goldwasser-Micali cryptosystem is performed by taking the product of two ciphertexts. This is a binary XOR.

Let us assume that we have two ciphertexts c1 and c2 which are respectively the encoded counterparts of plain text m1 and m2; the encryption of m1.m2 can be calculated by choosing
two random non-singular matrices:

\[ A, B \in (Z_2)^{l \times l} \]

Combined cryptotext, \( c = Ac_1 + Bc_2 \) [12] [13]

The homomorphic AND operation can be performed by a series of homomorphic XOR operations on the encrypted vectors which are given as input to the algorithm.

**D. The Boneh-Goh-Nissim Cryptosystem (BGN) – SWHE**

The cryptosystem devised by Boneh, Goh, and Nissim [14] was the first to allow both additions and multiplications with a constant-size ciphertext. There is a catch, however: while the additive property is the same as for the ElGamal variant, only one multiplication is permitted. The system is thus called “somewhat homomorphic.”

### 1. Key Generation

**Key Generation**

Step 1: select two large primes at random - \( p, q \)

Step 2: Find a supersingular elliptic curve \( E/\mathbb{F}_p \) with a point \( P \) of order \( n \)

\( W = \langle P \rangle \)

Step 3: Choose \( Q' \leftarrow W \setminus \{ \infty \} \) and set \( Q = [r]Q' \); then \( Q \) has order \( q \). Let \( e : W \times W \rightarrow \mu_n \subset \mathbb{F}_p^* \).

Step 4: Output the public key \( pk = (E, e, n, P, Q) \) and the private key \( sk = q \).

### 2. Encryption

**Encryption**

Input message \( M \in \mathbb{Z}_p \)

Step 1: generate a random number \( r \)

Step 2: compute: \( C = m + r x p^k \mod n \)

### 3. Decryption

**Decryption**

Input message is \( C \)

Step 1: calculate: \( m = c \mod p \)

Output: \( m \in \mathbb{Z}_p \)

Figure 5: EHES Algorithm

Let \( a, b \in \mathbb{Z}_p, pk = (n) \) and \( sk = (p, q) \)

Multiplicative: \( EncEH (a \times b) \equiv (EncEH (a) \times EncEH (b)) \)

Additive: \( EncEH (a + b) \equiv EncEH (a) + EncEH (b) \)

Where \( a \times b = DecEH (EncEH (a) \times EncEH (b)) \)

**E. Enhanced Homomorphic Encryption Scheme Cryptosystem (EHES) – FHE**

Enhanced homomorphic Encryption Scheme (EHES) for homomorphic encryption as well as decryption was posited by Gorti Vaikuntanathan Subba Rao, supplemented by research from Garimella Uma in 2013. This was carried out with the IND-CCA secure system. This Cryptosystem presents operations like addition, multiplication individually and mixed as well [16].

1. **Key Generation**

   Step 1: select two large primes at random - \( p, q \)

   Step 2: computing their system modulus \( N = p \cdot q \)

   Output is \( (pk, sk) \)

   Where \( pk = (n) \) \{public key\}

   And \( sk = (p, q) \) \{private key\}

2. **Encryption**

   Input message \( M \in \mathbb{Z}_p \)

   Step 1: generate a random number \( r \)

   Step 2: compute: \( C = m + r x p^k \mod n \)

3. **Decryption**

   Input message is \( C \)

   Step 1: calculate: \( m = c \mod p \)

   Output: \( m \in \mathbb{Z}_p \)

**F. Algebra Homomorphic Encryption Scheme based on Updated El Gamal (AHEE) – FHE**

The Algebra Homomorphic Encryption Scheme is a modified version of the DSS i.e., Digital Signature Standard proposed by NIST [17]. This algorithm is highly secure and can reliably maintain data confidentiality against plain text attacks and similar threats, thus resulting in varied applications. The AHEE can be referred to as subset of the fully homomorphism as it allows addition as well
multiplication to be performed on encrypted data without any cap limits on the number of operations performed.

<table>
<thead>
<tr>
<th>1. Key Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: select any two prime numbers - p and q, preferably large</td>
</tr>
<tr>
<td>Step 2: computing their system modulus N= p.q</td>
</tr>
<tr>
<td>Step 3: select random number h</td>
</tr>
<tr>
<td>Select a root r of GF(p)</td>
</tr>
<tr>
<td>Step 4: calculate y = r^h mod p.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Encryption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input message M is made up of multiple bits m1,m2…m3</td>
</tr>
<tr>
<td>Step 1 Select random integer number I</td>
</tr>
<tr>
<td>E1 (M) = (M+i* p) mod N.</td>
</tr>
<tr>
<td>Step 2: Choose a random integer k such that:</td>
</tr>
<tr>
<td>Er (M) = (a,b) = (rk mod p, yk E1 (M)mod p)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Decryption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input message C is made up of multiple bits c1,c2…c3</td>
</tr>
<tr>
<td>Step 1: M = b× (a^b) -1 (mod p)</td>
</tr>
<tr>
<td>Output message is M = m1,m2,…m3</td>
</tr>
</tbody>
</table>

Figure 6: AHEE Algorithm

Consider two messages M1 and M2 to be stored in an insecure environment:

Multiplicative:

EnAH (M1M2) = EnAH (M1)·EnAH (M2), or M1·M2=DeAH (EnAH (M1)·EnAH (M2)).

Additive:

EnAH (M1+M2) = EnAH (M1) ⊕ EnAH (M2), or M1+M2=DeAH (EnAH (M1) ⊕ EnAH (M2)).

These equations demonstrate additive and multiplicative properties of Homomorphic Encryption as seen in Algebra Homomorphic Encryption Scheme Based on Updated El Gamal (AHEE).

IV. COMPARATIVE STUDY OF THE EXISTING TECHNIQUES

As we have seen in the previous section, security is an important concern of the cloud computing environment. There are multiple techniques belonging to various categories that have been introduced to provide confidentiality. However, a fine balance between feasibility, speed and confidentiality is desired. Since users are looking for techniques that will provide the optimum results for them, we compare the techniques based on multiple relevant parameters – the category, the type of operations that can be performed and the possible applications based on this data to assist with the decision making process.

<table>
<thead>
<tr>
<th>HE Technique</th>
<th>Category</th>
<th>Operation possible</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSA</td>
<td>Partial HE</td>
<td>Multiplicative</td>
</tr>
<tr>
<td>GM Cryptosystem</td>
<td>Partial HE</td>
<td>Additive</td>
</tr>
<tr>
<td>SYY</td>
<td>Somewhat HE</td>
<td>AND Operation</td>
</tr>
<tr>
<td>BGN</td>
<td>Somewhat HE</td>
<td>Unlimited additions, one multiplication</td>
</tr>
<tr>
<td>EHES</td>
<td>Fully HE</td>
<td>Mixed</td>
</tr>
<tr>
<td>AHEE</td>
<td>Fully HE</td>
<td>Mixed</td>
</tr>
</tbody>
</table>

Figure 7: Categorical classification of Algorithms

<table>
<thead>
<tr>
<th>HE Technique</th>
<th>Suggested real life application</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSA</td>
<td>Credit Card transactions and Net Banking applications</td>
</tr>
<tr>
<td>GM Cryptosystem</td>
<td>For confidential E-voting and auctions, biometric-based authentication. [21]</td>
</tr>
<tr>
<td>SYY</td>
<td>Mobile Ad hoc networks for basic computations</td>
</tr>
<tr>
<td>BGN</td>
<td>Multiparty restricted computation</td>
</tr>
<tr>
<td>EHES</td>
<td>Consumer privacy in advertising, Financial Privacy [18]</td>
</tr>
<tr>
<td>AHEE</td>
<td>Medical applications, Data mining, Forensic image recognition [18]</td>
</tr>
</tbody>
</table>

Figure 8: Possible applications

V. FUTURE WORKS

When it comes to making a decision about the application of a particular Homomorphic encryption technique, we need to take into several factors such as the operations possible, the overhead involved in terms of time and space complexity and the type data to be encrypted. To combine the advantages observed in two standalone techniques, we have come across research that aims to create a hybrid Homomorphic Encryption algorithm [19]. As the research goes on to explain how RSA and El Gamal techniques are used in tandem to provide enhanced data confidentiality thereby augmenting the applications of such techniques. To provide a brief idea of this research, the algorithm used is such that plain text data is encrypted with RSA at the outset. This encrypted data is treated as the input data to a system which further encrypts it again with El Gamal. Computations are performed on this data and the decryption process follows the exact same process in the reverse manner.

We believe that, combining existing techniques in such a manner as to reduce the shortcomings while enhancing the advantages provided individually is promising area to carry out research. In our future works, we propose to compare the results of performing data mining on information stored on a system built using OwnCloud [20] which is encrypted using a hybrid algorithm with that of existing fully homomorphic encryption techniques like AHEE and EHES.

VI. CONCLUSION

Cloud computing offers individuals and businesses- small or large a wide range of services as mentioned above.
However, potential users are still hesitant when it comes to performing sensitive operations using cloud based technology. Most of this hesitation stems from lack of data confidentiality due to possible data leak. With this paper, we provide the reader with a basic understanding of Homomorphic encryption which proves to be an outstanding solution to the issues regarding data confidentiality. We also provide a categorization of the various Homomorphic encryptions based schemes and provide examples of each category. We further provide the reader with an overview of the operations possible with these algorithms and the real life suggested applications of the studied algorithms. The comparisons provided and the possibility of future research mentioned is done so with the aim of providing assistance to extend the existing and on-going research in this field to improve on data confidentiality in cloud based computing.

VII. REFERENCES


[18]. Fredrick Armknecht et al., “A guide to fully Homomorphic Encryption”,2015, pp 6-10


