# Data Protection for Users in Cloud Computing

Augustine Obulor Ugbari
University of Port Harcourt
Computer Science Department
Faculty of Science

Betty Osamegbe Ahubele
Benson Idahosa University
Computer Science Department
Faculty of Science

## ABSTRACT

Information technology might have simplified the way we do business, but it has also brought challenges that never existed. Accessing data from public repositories leads to multiple difficulties. It is also necessary for the data to be available in an accurate, complete, and timely manner. This paper will discuss the challenges and breaches of data protection for users in cloud computing and explore various techniques used to address these issues.

## General Terms

Cloud, Security

## Keywords

Data protection, encryption techniques, cloud computing

## 1. INTRODUCTION

Cloud computing is an advanced method of accessing resources and services from a physical computer, therefore it doesn’t require a physical computer unlike the 80s and early 90s where users would be limited to resources provided by servers and host computers in an office or a building; a good example of cloud computing is computing based on the internet.

Cloud computing can be defined in so many ways, but Glossary defined it as an act of computing where expandable and flexible online resources are available as a service to every consumer making use of the internet.

Cloud computing is therefore the access to unlimited resources that are propertied and managed by a third party (host computer) through the internet. A host computer is a computer that holds resources and stays on 24 hours a day to always give back to the users the resources it holds for the users.

The cloud is made of networks, storage, hardware, and interfaces that provide services aimed at preferring the accessibility in which users can get data irrespective of their locations. Users are liable to get this service with the aid of internet connections by paying for them. Every user uses cloud computing in one way or the other. For instance, with various social media platforms, you can access the services of the application regardless of time or location. The key point of cloud computing is that it enables the user to access resources without being present with the cloud components. Furthermore, cloud allows for data storage and data accessibility for its users leaving the hardware and software maintenance and data security for the administrators.

As asserted by NIST, Cloud computing is the means for easy access to a collective network of resources that can be immediately gotten and sent out with little assistance or help from service providers.

While cloud computing tackles many modern-day computing problems, it has also engendered new challenges mainly on data security. Data security problems might be tackled by cryptographic algorithms since their function is to hide data from unauthorized users.

For cloud data storage to be possible, users need to send their data to an administrator who will maintain and store their data.

Encryption algorithm plays a key role in data security under cloud computing; encryption is the transforming of data into unreadable form during storage and transmission that it appears waste to intruders while decryption is basically the reverse of encryption.

There are three main encryption methods: symmetric, asymmetric, and hashing algorithms. Symmetric algorithms encrypts and decrypts with only one key, while asymmetric algorithms encrypts and decrypts with two keys.

## 2. MERITS OF CLOUD COMPUTING

There are various business/commercial benefits to cloud computing. It allows you to set up what is essentially a virtual office to give you the flexibility of connecting to your business anywhere, any time. With the growing number of web-enabled devices used in today's business environment (e.g., smartphones, tablets), access to your data is even easier. Key benefits are outlined and discussed below:

### Flexibility

Cloud computing helps businesses meet demands for their customers by providing access to more resources when necessary - by creating high-capacity servers, storages, and other services, it makes it easy to meet any strategic time needs. This feature allows users to achieve their aim notwithstanding the project size. Businesses that have fluctuating bandwidth needs are perfect for cloud-based services because of theirscalability aimed at providing more cloud capacity if needed. This flexibility feature provides business an edge over other competitors in cloud computing.

### Reduced Capital Expenditure

Cloud computing provides a subscription platform for users, it’s basically a pay on demand service that is economical – all this is aimed at reducing increased cost of hardware. Cloud computing services are accessible to various pools of users in the world, its billing system depends on the level of usage of infrastructure and other services which also enables users reduce cost by being more specific in what they want. The service demands of businesses can also be scalable based on their performance level in the market.

### Controllability

Using central administration of resources, vendor managed infrastructure and Service Level Agreement (SLA) backed agreement, cloud computing provides simplified and better information technology maintenance and manageability. All resources are controlled by service providers which allows users access a simplified user interface that is web-based for locating and using applications, software's and other services that is all carried out without installation, and in which the SLA guarantees on-time delivery, and adequate maintenance of your IT services.

**Dependability**

Comparatively, cloud computing is more efficient and dependable than hardware-based IT infrastructure due to its manageable/tamable service platform. Furthermore, a high number of cloud-based service providers provide an SLA that allows service availability anytime. Businesses and users enjoy more IT resources and well as an efficient fail-save model that is, in case a server malfunctions, there is easy transferabilityof those services and applications to another available servers.

**Remote Work**

Cloud computing provides users with the luxury to be at work anywhere if a viable internet connection is available. There are no work-related restrictions, and this provides more flexibility to businesses in the sense that workers can have an ideal balance to their work-life without jeopardizing their productivity level.

**Data Protection**

Loss of data is a most dreaded situation for businesses. With the advent of cloud computing, there is increased security since there is cloud-based storage of data. This makes it possible for users to get access to their data irrespective of what happens to their devices and users can also decide to delete their data from lost devices remotely if it runs the risk of being accessed by intruders. Nonetheless, data security also poses a challenge in cloud computing.

**Regular Automated Updates**

Frequent automated updates are one of the key advantages of cloud computing that is, the servers are out of sight and service providers can frequently update the software and security so that users enjoy better service experience without worrying about system maintenance issues.

**Multi-sharing**

With the aid of multi-tenancy and virtualization features, cloud computing enables sharing of resources through the internet among various users. Furthermore, with cloud computing various users can access resources and perform tasks efficiently at reduced cost due to sharing common data/service infrastructures.

# 3. CLOUD SECURITY ISSUES

## 3.1 Normal or Body Text

Security is an immense problem when remote users store vital information onto a platform that is not managed and controlled by themselves. Data security is key because during the sending and storage of data, data can be unauthorizedly assessed and modified by intruders. Data is termed secure if it meets these three criteria: (i) it is confidential, (ii) Data integrity, and (iii) data availability.

Data is confidential only when the data is eligible to the receiver alone and no other person - this prevents unauthorizedaccess to the user's sensitive information.

Data integrity implies that the data is received in the exact form and structure as the sender sent it. This ensures that the data is not tampered with by unauthorized users.

Data availability simply means that users can access resources irrespective of time and location. Cryptography is the method through which data confidentiality is achieved. There are variegated measures for ensuring data security like proffering encryption methods and access controls. It is the task of the service providers to guarantee that the service infrastructures and clientele's data are secured. Users should also make sure that their data is secured on their own end as much as the administrators are providing security measures for them.

## 3.2 PROBLEM FORMULATIONS

Cloud computing is a large depot of networks that are linked together. Like any data related technology, there are threats and policy issues such as storage, privacy, data reliability, data integrity, data confidentiality, availability and more. Since sensitive information/data are stored in the cloud, data security becomes one of the key critical issues in modern day cloud computing. Furthermore, there are innovative services and technologies that have not been vetted in line with security. So, our focus is on data security of cloud computing.

# 4. EXISTING ALGORITHM

Users store their data on cloud for other authorized users to access, therefore we will need to protect these data from unauthorized users. To ensure data security for many businesses and users, various algorithms are created. The very effective and recommended algorithms are outlined and discussed below.

## 4.1 Data Encryption Standard(DES)

DES is a symmetric key-block cipher brought about by NIST (National Institute of Standard and Technology) in 1977. Since it is symmetric, it encrypts and decrypts with a single key while also operating on 64-bits blocks of data with 56 bits keys and round key size of 48bits.

The functionality of the block-cipher is to encrypt data blocks which have simple text by both confusion and diffusion. To achieve cipher block it must go through 16 rounds, the 64bits must be split equally before going through the 16 rounds. After splitting the 64bits, the Feistel function (F-function) is applied. The Feistel function comprises of permutation, key mixing, substitution. The result of the function is added with the other half data making use of XOR gate therefore data crossing is done. When 16 rounds have been done cipher text will be produced, that is the data has been successfully encrypted. To gain access to this encrypted data, the encryption process must be reversed. One big limitation of DES is that intruders can easily bypassits security due to it using very small keys for encryption. Another limitation could be it has slow response on software but not hardware.

## 4.2 Advanced Encryption Standard (AES)

4.3 AES is popularly known as Rijndael. It is also a cipher symmetric key block published by NIST in 2000. AES contains various sizes of key i.e., 128,192 or 256 bits, depending on its rounds. 14 rounds use 256 bits key, 12 rounds make use of 192 bit key and for 10 rounds 128-bit key is used. All cycles of AES are the same except the final round. cycles==round. AES makes use of a 4×4 matrixes. It also comprises of expansion key, initial cycles, and final cycles. The initial cycle contains sub bytes, mix column, shift rows, add round key while final cycle contains all function of the initial cycle but mix columns. Unlike DES, AES has a better

response to software and hardware.
AES is preferred over DES for the following reasons:

1. 128 bits is it data block size
2. there are various key sizes
3. permutation and substitution are used
4. hardware AES is included in most CPU, increasing its response time.
5. 2128,2192 and 2256 are its possible keys
6. it's the most frequently used symmetric encryption algorithm
7. it has high security level than that of DES

## 4.4 Triple Data Encryption Algorithm (Triple- DES)

Triple DES(TDES) is another symmetric key block cipher developed in 1988. It is like DES but applied thrice to increase security level. It can be said to be a better version of DES. Though 3DES is slower than other block cipher methods, TDES only enables the increase of size of key while the other works are like DES. 3DES is lower in performance than DES. Due to its triple phase encryption, due to its triple phase, it requires more time than DES.

## 4.5 Blowfish Algorithm

The blowfish algorithm was created by Bruce Schneier in 1993. It is a symmetric key algorithm that has a somewhat similar working operation to DES. The main difference being that the DES has a reduced key size and be decrypted easily whereas the blowfish algorithm has a pronounced key size ranging from 32 to 448 bits and comprises of 16 rounds like DES. more of the characteristics of Blowfish algorithm are that it can perform encryption on data having a size multiple of eight and can pad data if the size is not a multiple of eight, and it divides 64bits of plain text into 32 bits parts. Numerous research and experiments have asserted that Blowfish algorithm is superior to other algorithms when it comes to processing time and power consumption.

## 4.6 IDEA

James Ramsey Massey and Xuejia Lai in 1991 developed the International Data Encryption Algorithm which is generally affirmed to be the best symmetric key algorithm that accepts 64 bits plain text and a 128 bits key size. It is made up of 8.5 rounds which all look the same but the one. Furthermore, it splits 64 bits of data into 4 blocks where each block is 16 bits in size and the sub blocks have basic operations modular, multiplication, addition and bitwise exclusive (XOR) applied on them. It also has characteristics like; IDEA has 8.5 rounds each comprising of different sub keys where the total number carrying out different rounds are 52 sub keys.

In the first round, the K1 to K6 subkeys are created, in which the sub key K1 carries the foremost 16bits of the original key and K2 the next set of 16bits and so it applies to the K3 to K6. Ergo, 96 bits of cipher is used for the first round, that is, (16*6 = 96). Outlined below is the sequence of operations carried out in each round. Where 11, 12, ...16 are the inputs to round 1, therefore the functions in round 1 are:

i. Multiply I1 and K1.
ii. Add I2 and K2.
iii. Add I3 and K3.
iv. Multiply I4 and K4.
v. Now, step 1 is EXOR with step 3.
vi. Step 2 EXOR with step 4.
vii. Multiply step 5 with K5.

The same operations are carried out in each respective round.

## 4.7 Diffie- Hellman Key Exchange

In 1976, Whitfield Diffie and Martin Hellman created the Diffie Hellman key exchange algorithm, and it is one of the foremost workable examples of key exchange in cryptography. Its method of exchanging cryptographic keys is specific and permits two unknown parties to create a common security key knowable only to them in an insecure channel of communication. This common key might also be used subsequently to encrypt other communications. One key thing in the Diffie Hellman Key Exchange Algorithm is that both sender and receiver can select two secret digits and these digits are only knowable to both sender and receiver. Below is how the cryptographic algorithm works.

Suppose Alice (A) and Bob (B) want to agree upon a key which they will use to encrypt their subsequent messages to one another. The key exchange between A and B works as follows:

1. A and B agree on a finite field Fq and a generator g of the cyclic group F∗q

2. A chooses a random integer a between 1 and q−1, which she keeps secret, and computes ga ∈ Fq, which she sends to B.

3. B chooses a random b, computes gb ∈ Fq and sends gb to A. He keeps b secret.

4. A computes KA =(gb)a

5. B computes KB = (ga)b

KA should be equal to KB and this key is only known to A and B. The secret key they use is then gab. A and B can now use this private key to communicate using some cryptographically secure communication protocol. During this key exchange, the values ga and gb are publicly known. It is computationally infeasible to compute gab from the known values ga and gb. If the discrete logarithms were easy to compute, then one would compute a = logg(ga) and then compute (gb)a. It is not known if there is an easy way to compute gab from knowledge of ga and gb without computing the discrete logs a and b.

## 4.8 ElGamal

This is another public key cryptography presented by Tahar ElGamal in 1984. ElGamal simplifies the Diffie-Hellman key exchange which requires interaction of both the sender and receiver to calculate a common private key. This makes it extremely difficult especially when the cryptosystem is applied to communication channels where both parties are not able to interact due to several factors. Hence, ElGamal introduced a random exponent k. Below illustration shows how ElGamal works:

1. Fix a very large finite field Fq and an element g ∈ Fq (preferably a generator).
2. Randomly select an integer b in range 0 < b < q − 1. This is the secret deciphering key.
3. He then computes gb and makes public (q, g, g) b. This is Bob's public key.

Assuming we are using plaintext message units with numerical equivalents m in Fq, Alice encrypts the message as follows:

1. 1. Obtain Bob's public key (q, g, g) b.
2. 2. select a random exponent k and compute gk ∈ Fq and m(gb)k
3. 3. Send Bob the pair of elements of Fq; (gk, mgbk)

## 4.9 Homomorphic Encryption

This encryption makes use of asymmetric key algortihms where two separate keys are used for both encryption and decryption (these two keys are termed public key and private key). Mathematically, homomorphic is defined as the conversion or transforming of one data set to another, with the relationship between them still intact. In this kind of encryption, complex mathematical functions are used for encryption of data and the vice versa of the encryption operation is applied for the decryption of data

## 4.10 RSA

Ron Rivest, Adi Sharmir, and Leonard Adleman in1977 created the RSA as an internet authentication system which uses an algorithm. Even today, RSA is one of the frequently used algorithms which is used for the generation and encryption of both public and private keys. One of its key characteristics is that its encryption operation is fast while being an asymmetric algorithm. Its functionality is founded on the multiplication of two large prime numbers, which afterwards modulus is computed, and the resultant number is used as the private and public key. Furthermore, the two numbers in which the multiplications are performed on arepublic while the other is private. Outlined below are the steps for RSA algorithm:

i) the large message is split into smaller blocks in which each block stands for the same range

ii) Raise the eth power to module and encrypt the message

iii) when carrying out decryption of the message, increase another power d module n.

RSA Cryptosystem involves two aspects: Generation of Key Pairs and Encryption-Decryption algorithms. Generation of key pairs must involve each person or party who desires to participate in communication using encryption process to generate a pair of keys, namely public key, and private key.

RSA Key Generation Algorithm:

Output: public key: kpub = (n, e) and private key kpr = d, where n=RSA Modulus (a minimum of 512 bits), Kpub=Public Key, kpr= Private Key and e=number greater than 1 and less than (p-1)(q-1).

1. Choose two large primes p, q
2. Compute n = p * q
4. Compute Φ(n) = (p-1) * (q-1)
5. Select the public exponent e ε {1, 2, …, Φ(n)-1} such thatgcd(e, Φ(n)) = 1
6. Compute the private key d such that d * e ≡ 1 mod Φ(n)
7. RETURN kpub = (n, e), kpr = d

**RSA Encryption and Decryption:**

Successful generation of key pairs makes the process of encryption and decryption straightforward and computationally easy. However, RSA does not directly operate on strings of bits as in case of symmetric key encryption but on numbers modulo n. Consequently, the plaintext is represented as a series of numbers less than n. The Encryption-Decryption algorithm is defined as follow:

Given the public key (n,e) = kpub and the private key, d = kpr; we write

y = ekpub(x) ≡ xe mod n

x = dkpr(y) ≡ yd mod n

where x, y ε Zn, we call ekpub() the encryption and dkpr() the decryption operation.

Practically, x, y, n and d are very long integer numbers (≥ 1024 bits). Hence. the security of RSA relies on the fact that it is hard to derive the "private exponent" d, given the public-key (n, e). Since RSA has more functionality, while AES is much faster, it is best to combine these two algorithms. Asymmetric cryptograph can be used to authenticate the parties and to agree on a key for symmetric encryption and large data blocks are encrypted using faster AES algorithm rather than slower RSA and safely distributed using RSA algorithm.

## 4.11 ECC (Elliptic Curve Cryptography)

ECC is more efficient than RSA. RSA and Elgamal asymmetric cryptography require exponentiations in integer rings and fields with parameters greater than 1000 bits. In practice, ECC offers high computational effort on CPUs with 32-bit or 64-bit arithmetic and large parameter sizes critical for storage on small and embedded systems. The need for smaller field sizes providing equivalent security became necessary. Hence the invention of ECC. Elliptic Curve Cryptography uses a group of points instead of integer module for cryptography with coefficient sizes of 160-256 bits thereby significantly reducing the computational effort. ECC uses complex mathematical algorithms for data protection. ECC offered higher security level which makes it more valuable for digital signatures in cryptocurrencies such as Bitcoin and Ethereum in signing transactions.

## 5. CONCLUSION AND FUTURE SCOPE

Conclusively, it is now believed generally that cloud computing is a blooming and technology with lots of promise and applicability to the next generation of IT applications and organizations. Due to the voluminous data in organizations, cloud proffers a space where accessibility of data can be performed irrespective of time and place. Data security is the most renown problem in cloud computing as virtualization is one of its key merits. With dependable data security, cloud computing can be a very useful technology worldwide since it allows for storage of large volumes of data to its users. Various algorithms are now available for cloud computing, each with its own method of operations. These algortihms such as the DES, AES, and the Triple DES are termed symmetric key algorithms, that is, the perform encryption and description with the use of a single key while RSA, Diffie-Hellman Key Exchange and Homomorphic equations are termed asymmetric, that is, encryption and decryption are performed with the use of two different keys.

## 6. REFERENCES

[1] Alexa Huth and James Cebula 'The Basics of Cloud Computing', United States Computer Emergency Readiness Team. (2011)

[2] Garima Saini, Gurgaon Naveen Sharma,” Triple Security of Data in Cloud Computing “, Garima Saini et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol.5 (4), 2014.

[3] D. S. Abdul. Elminaam, H. M. Abdul Kader and M. M. Hadhoud, “Performance Evaluation of Symmetric Encryption Algorithms”, Communications of the IBIMA Volume 8, 2009.

[4] Qi. Zhang ·Lu. Cheng, Raouf Boutaba, “Cloud computing: state-Of-the-art and research Challenges”, “The Brazilian Computer Society”, April 2010.

[5] Foster, I. T., Zhao, Y., Raicu, I., & Lu, S. (2009). Cloud Computing and Grid Computing 360- Degree Compared CoRR. abs/0901.0131.

[6] Srinivasa rao v, Nageswara rao n k, E Kusuma kumari, “Cloud Computing: An Overview”, Journal of Theoretical and Applied Information Technology.

[7] Rachna Jain and Ankur Aggarwal ‘Cloud Computing Security Algorithm’, International Journal of Advanced Research in Computer Science and Software Engineering. January (2014) Vol. 4, Issue 1.

[8] Manzoor Hussain Dar, Pardeep Mittal, and Vinod Kumar, ‘A Comparative Study of Cryptographic Algorithms’, International Journal of Computer Science and Network. June (2014) ISSN(Online): 2277- 5420, Volume 3, Issue 3.

[9] Gartener: Seven cloud-computing security risks. InfoWorld.2008-07-02. http://www.infoworld.com/d/security-central/gartener-seven-cloud- computing-security-risks-853.

[10] G. Devi and M. Pramod Kumar, ‘Cloud Computing: A CRM Service Based on a Separate Encryption and Decryption using Blowfish Algorithm’, International Journal of Computer Trends and Technology. (2012) Vol. 3 Issue 4, ISSN: 2231-2803, pp.592-596.

[11] Mr. Gurjeevan Singh, Mr. Ashwani Singla and Mr. K S Sandha “Cryptography Algorithm Comparison for Security Enhancement in Wireless Intrusion Detection System” International Journal of Multidisciplinary Research Vol.1 Issue 4, August 2011.

[12] Sandipan Basu, ‘International Data Encryption Algorithm (IDEA) - A Typical Illustration’, Journal of Global Research in Computer Science. July (2011) ISSN: 2229-371X Vol. 2, Issue 7.

[13] Ayan Mahalanobis, ‘Diffie-Hellman Key Exchange Protocol’, Its Gernalization and Nilpotent Groups. August (2005).

[14] Maha TEBAA, Said EL HAJJI and Abdellatif EL GHAJI, ‘Homomorphic Encryption Applied to the Cloud Computing Security’, World Congress on Engineering. July 4 (2012) Vol. 1, London U.K. ISBN: 978-988-19251-3-8, ISSN: 2078-0958 (Print); ISSN: 2078-0966 (online).

[15] B.Persis Urbana Ivy, Purshotam Mandiwa and Mukesh Kumar, ‘A Modified RSA Cryptosystem Based on ‘n’ Prime Number’, International Journal of Engineering and Computer Science. Nov (2012) ISSN: 2319-7242 Volume 1 Issue 2.

[16] Shakeeba S. Khan and Prof. R.R. Tuteja, ‘Security in Cloud Computin Using Cryptographic Algorithms’, International Journal of Innovative Research in Computer and Communication Engineering. January 1, (2015) ISSN (online): 2320-9801, (Print): 2320-9798 Vol. 3, Issue.

[17] Ohri, A. (2020). Elliptic Curve Cryptography: An Overview. Retrieved 27/01/2022www.jigsawacademy.com

[18] Muyinnda, N (2014. Elliptic Curve Cryptography. Mekere University. Retrieved 28/01/2022. www.researchgate.net

[19] Gahan, A. (2019). An Empirical Study of Security Issues in Encryption Techniques. International Journal of Applied Engineering Research. 14(5). 1049-1061