Privacy Enforcement on Subscribers Data in Cloud Computing

S. A. Akinboro\textsuperscript{a,*}, U. J. Asanga\textsuperscript{b}, M. O. Abass\textsuperscript{a}

\textsuperscript{a}Department of Computer Science, University of Lagos, Akoka, Lagos State, NIGERIA.
\textsuperscript{b}Department of Computer Science and Information Technology, Bells University of Technology, Ota, Ogun State NIGERIA.

Abstract

Data stored in the cloud are susceptible to an array of threats from hackers. This is because threats, hackers and unauthorized access are not supported by the cloud service providers as implied. This study improves user privacy in the cloud system, using privacy with non-trusted provider (PNTP) on software and platform as a service model. The subscribers encrypt the data using user’s personal Advanced Encryption Standard (AES) symmetric key algorithm and send the encrypted data to the storage pool of the Cloud Service Provider (CSP) via a secure socket layer. The AES performs a second encryption on the data sent to the cloud and generates for the subscriber a key that will be used for decryption of previously stored data. The encryption and decryption keys are managed by the key server and have been hardcoded into the PNTP system. The model was simulated using the Stanford University multimedia dataset and benchmarked with a Privacy with Trusted cloud Provider (PTP) model using encryption time, decryption time and efficiency (brute force hacking) as parameters. Results showed that it took a longer time to access the user files in PNTP than in the PTP system. The brute force hacking took a longer time (almost double) to access data stored on the PNTP system. This will give subscribers a high level of control over their data and increase the adoption of cloud computing by businesses and organizations with highly sensitive information.

Keywords: privacy, AES, subscriber, CSP, PTP, PNTP, brute force hacking

1. INTRODUCTION

Cloud computing is a computing model that enables convenient on demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management efforts or service provider interactions [1]. The main aim of cloud computing is to reduce the need for customers’ investment in new hardware or software by offering flexible cloud services, with a subscriber reaping the benefits of the pay per use approach. Cloud computing demands addressing security and privacy issues like vulnerabilities, threats, and attacks and proffering possible solutions. Threats in cloud computing may include data breaches, human error such as accidental deletion of data by the cloud service provider or physical catastrophe, malicious insider, account hijacking and distributed denial of service. Identified vulnerabilities in cloud computing are consumers having reduced visibility and control, on-demand self-service which has simplified unauthorised use of cloud services, Internet-accessible management APIs been compromised, failing of separation among multiple tenants, and incomplete data deletion.

A classic definition of security in terms of its basic characteristics is confidentiality, integrity and availability which are the three key requirements for any secure system [1]. Confidentiality is the ability to hide information from those people unauthorized to view it. It is the basis of many security mechanisms protecting not only information but other resources. Integrity is the ability to ensure that the data are accurate and unchanged representation of the original information. Availability ensures that a resource is readily accessible to the authorized subscriber upon the subscriber’s request [1].

Privacy is the right to have information about oneself left alone [2] or the selective control of access where individuals control their interaction and information exchange with others [3]. To assure their privacy, individuals try to control their openness to others based on their relationship and the value given to the information [4]. Despite cloud computing’s widespread acceptance, security and privacy issues resulting from the illegal and unethical use of information or disclosure of confidential information has hin-
dered businesses from adopting cloud-based services due to misconduct that can be performed by the service provider. Subscribers to cloud-based services need to be assured of confidentiality, integrity and availability of their data to gain their confidence on the use of the platform. As a result, numerous researchers have studied and surveyed the issues of security and privacy in cloud environments over the years. To induce subscriber’s confidence, there is need for an efficient system which performs authentication, verification and encrypted data transfer. While cloud computing is associated with numerous security and privacy problems, it can be improved by implementing efficacious solutions.

The study separated cloud computing security issues from its privacy issues while confidentiality, integrity and availability are ensured by proposing a Privacy with Non-Trust Provider (PNTP) system. The subscriber’s data are categorized based on its level of sensitivity. The system will give its subscriber the right to hide very sensitive information from unauthorized persons including the cloud service providers.

The rest of the paper is arranged as follows: Section 2 gives some of the existing work; Section 3 discusses description of the proposed system, algorithm for secure storage and accessing data from CSP or PNTP category; system implementation, result and analysis are given in Section 4; conclusion and future work are drawn in Section 6.

2. LITERATURE REVIEW

The main aim of cloud computing is to reduce the need for customers’ investment in new hardware or software by offering flexible cloud services, with a subscriber reaping the benefits of the pay per use approach. Customers’ concern about privacy issues remain a major barrier for the adoption of cloud computing services and platforms [5–7].

Cloud subscribers do not have access to the cloud’s internal operational details; therefore Cloud Service Providers (CSP) may voluntarily examine subscribers’ data for various reasons without detection [8]. An example is the Kenyan election of September 2017 which was cancelled due to information tampering and the 2012 attack to the cloud where 50 million subscriber accounts of Drop box were hacked [9]. [10] makes an argument for identity management system to achieve more automatic and fast subscriber account provisioning and de-provisioning in order to ensure no un-authorized access to organizations’ cloud resources by some employees who have left the organizations since mobility of employees in some organizations is relatively high.

To secure customers data entrusted to cloud servers’ owner, [11] proposed an authorized deduplication scheme using Ciphertext-Policy Attribute Based Encryption (CP-ABE). The scheme solved the problem of storage of data, data sharing and control over access permissions in encrypted deduplication storage. The scheme provided client-side deduplication while providing confidentiality to prevent exposure of subscribers’ sensitive data on untrusted cloud servers. The system provided an adequate trade-off between storage space efficiency and security in cloud environment which is suitable for hybrid cloud model. Loss and manipulation of data from unknown sources was prevented by providing secure computing environment, which is defined as a system implemented to control storage and use of data [12].

[13] provided an enhanced technique for improving security and protecting the privacy of the cloud computing subscribers by encrypting the data before it reaches the server’s storage. Analysis of cloud computing issues on data integrity, privacy and its current solutions were discussed by [14]. Integrity check was done by the data owner or a third party auditing by using RSA and MD5 (Rivest-Shamir-Adleman and Message Digest algorithm 5) cryptographic algorithm to avoid overwhelming workload for the data owner. [15] developed an authorization system to implement network security and protect sensitive data of each patient. The system is an access control model in hybrid cloud for healthcare systems, which handled security vulnerabilities. It includes protecting share data containing basic information of patients among hospitals in the system and protecting private patient data that can be accessed only by the treating doctor. The model was implemented in a real-case application to demonstrate its effectiveness in managing different levels of security and privacy. [16] proposed the security of big data in the cloud, such that only legitimate subscribers will have access to the cloud services. Their framework considers a data owner encrypts data with clock timing when storing data to the cloud storage. When subscribers try to access the data, it is done by permission according to its task and role. The research was able to achieve full secured and authorized access to cloud with big data. [17] identified challenges in maintaining multimedia data security and privacy for mobile cloud subscribers. They proposed image encryption technique called privacy preserving lightweight image encryption (PPLiIE). The PPLiIE algorithm proceeds with a three step process to secure the image data in mobile before storing to the cloud. The encryption time of PPLiIE was reduced by 50% approximately than the encryption time of AES algorithm. The measurement of key sensitivity and file with variation of chunk size expressed superior performance of PPLiIE. [18] discovered new privacy challenges originating in emerging new usage requests, on the accumulated content from multiple sources of various integrated devices at the Edge. Privacy content of multiple sources was modelled as resources of types of Data, Information and Knowledge known as (data, information, knowledge, wisdom) DIKW architecture. They categorized content objects and relationships uniformly as typed resources of DIKW comprising of Meta model of DIKW and extended data graph, infor-
Various cryptography techniques have been proposed to handle challenges in cloud computing to provide a secured computing environment where data confidentiality can be maintained [1, 22–25]. [26] Designed a concrete privacy preservation incentive and rewarding (PPIR) scheme using bilinear pairing and group oriented cryptography technique. This was proved in the random model. The PPIR scheme was proven secure and efficient using communication cost and computational cost as performance metrics. [27] utilized cryptography and access control to ensure confidentiality, integrity and proper control over sensitive data. They designed their model using an enhanced RSA encryption algorithm and a combination of role-based access control model with Extensive Access Control Mark-up Language (XACML). The RSA encryption algorithm was used to secure the data in the cloud, while data access was through access control model with encryption and decryption having minimum time and cost.

To induce subscriber’s confidence, there is need for an efficient system that can perform authentication, verification and encrypted data transfer. While cloud computing is associated with numerous security and privacy problems, it can be improved by implementing effective solutions. The proposed privacy with non-trusted provider system will afford the subscriber an opportunity to take charge of access to highly sensitive data by performing a first encryption on the raw data before uploading and handing over the data management to the cloud service providers. However, intending user will not store data with potential government interest. Also data with potential government interest will not be encrypted and stored aboard the platform of the cloud service provider.

3. PROPOSED SYSTEM DESCRIPTION

The proposed system in Fig. 1 comprises of cloud subscriber and non-trusted cloud service provider. The framework is an infusion of subscriber’s security and cloud security on Privacy with Non-Trusted Provider (PNTP) system. The PNTP system consists of the subscriber encryption which allows the subscriber to encrypt their data by using subscriber’s personal symmetric key before the data is sent to the cloud service providers (CSP). The subscriber uploads encrypted data to the secure cloud and the Non-trusted provider (NTP) makes use of the Advanced Encryption Standard (AES) to perform a second encryption on the data. The NTP generates for the subscriber a key that will be used for decryption whenever the cloud data is retrieved. The encryption and decryption keys are managed by the key server and have been hardcoded into the proposed privacy with non-trusted provider system. The cloud encrypted data are sent to the cloud database for storage.

Algorithm for Secure Data Storage and Retrieval from PNTP

In this category, the data contains highly sensitive information that needs to be concealed from the CSP. The data is encrypted on subscriber’s side before uploading to the secure cloud. First, the data categorization mechanism is applied to check for highly sensitive data. Then the subscriber encrypts the data making use of AES registered in a trusted module and uploads the encrypted data to the storage pool of the CSP via secure socket layer (SSL). Data retrieval entails the subscriber logging into the CSP and sends a request for data access. The CSP checks the data category for sensitivity and asks the subscriber to register on the trusted module to check the validity of the registration details. If valid, trusted module informs the CSP and sends an encrypted symmetric key $k_2$ to the subscriber. The CSP sends data to the subscriber who makes
use of the symmetric key to decrypt the data. To access the original data, the subscriber further uses their personal decryption key $k_2$ from the first encryption. The algorithms were partitioned into algorithm 1 for data categorization on PNTP system, algorithm 2 to retrieve sensitive data from PNTP system and algorithm 3 for validation of data retrieval request. The flowchart for algorithm 1, algorithm 2 and algorithm 3 are shown in Figs. 2, 3 and 4 respectively.

**Algorithm 1: Data Categorization on PNTP System**

// category is set based on the sensitivity of data; highly sensitive data is encrypted before uploading to CSP
// Subscriber checks for data category:
1. IF category is set
Input: DPC = PNTP //here DPC represents data privacy category
   THEN do
   ii. DSD ← EKPrDO (EKPubCSP (PNTP, data id, owner id, EK (data,K1))) // subscriber encrypts sensitive data with AES algorithm before upload to CSP
   //DSD represents the highly sensitive data of the subscriber
   iii. DSD ← CSP // subscriber uploads encrypted data to CSP
   iv. REGDO ← EKPrDO (EKPubTM (PNTP, data id, owner id, K2)) //subscriber registers with trusted module
   v. K2 ← Generate symmetric key // CSP performs a second encryption using AES and generates symmetric key
   ELSE
   v. PRINT: (No privacy required)
   //subscriber data is not sensitive
   END IF

**Algorithm 2: To Retrieve Sensitive Data from PNTP system**

1. input: REQsub // Subscriber logs in to CSP and sends request to retrieve data
   ii. ACK ← REQsub //CSP acknowledges request
   iii. REGDO ← EKPrSub (EKPubTM (Uid, access control required, owner id, data id, PNTP)) //Subscriber registers with trusted module and provides k2
   // trusted module receives the registration information from subscriber and checks for validity
   iv. IF (DPC = PNTP) and K2 is valid // the data DPC is sensitive
   THEN
   v. Fetch Data ← (PNTP,data id, owner id, EK(data,k1)) // subscriber decrypts the data DPC with public key k2 and his own private key k1
   ELSE
   vi. PRINT: subscriber request cannot be granted
   END IF

**Algorithm 3: Validation of Data Retrieval Request**

1. Input: ACDO = ACrequested // request is valid if the requested access control is equal to access control of data permitted by subscriber
   ii. IF ( ACDO = ACrequested )
   THEN

Figure 1: Proposed System Design for the PNTP System.
Figure 2: Flowchart showing Data Categorization on PNTP System.
Figure 3: Flowchart to Retrieve Sensitive Data from PNTP system.
Figure 4: Flowchart showing Validation of Retrieval Request.

Figure 5: Home Page Interface.
iii. PRINT ‘Request is valid’
iv. // Trusted module sends subscriber information to CSP as:
REGDO ← EKPrTM (EKPubCSP (Uid, access control, owner id, data id))
// Trusted module grants permission to CSP to send the encrypted data EK (data, k2) to subscriber
v. Trusted module sends the symmetric key K2 to subscriber i.e. K2 ← EKPrTM (EKPub sub (K2)) // K2 represents encrypted symmetric key
vi. DSD ← CSP // CSP sends the encrypted data EK (data, k2) to subscriber via SSL.
vii. DSD ← K2 // Subscriber then decrypts encrypted data received from CSP using symmetric key K2.
viii. IF (Subscriber accesses the original data successfully)
THEN
ix. // Subscriber decrypts DSD using private key, K1 i.e.:
DSD ← K1 // subscriber uses k1 to perform a second decryption to access highly sensitive data
ELSE
x. PRINT: (Invalid Request )
END IF

4. SYSTEM IMPLEMENTATION

An AES 256 encryption and decryption key was hard coded into the PNTP system implementation using Microsoft C# Net. Software as a service (SaaS) and platform as a service (PaaS) cloud models were considered. The subscriber sends encrypted data to the CSP who performs a second encryption on the data and the time it takes for the encryption to be successfully completed is recorded by the system as one of the parameters for evaluation of the system processes. The encrypted data are then stored on the cloud database of the CSP secure cloud. At data retrieval, the CSP authenticates the request of the subscriber then checks the data category. The system provides the subscriber’s symmetric key managed by the key infrastructure and if it is correct, the CSP then performs a server decryption on the CSP encrypted data and returns to the subscriber, otherwise an error report is returned. The decryption time is also recorded by the system as evaluation parameter.

To measure the efficiency of the privacy with non-trusted provider (PNTP) over privacy with trusted provider (PTP), brute force hacking was applied on both systems. Assuming that a hacker has the correct encryption and decryption key used by the server, brute force then attacks the system to try and retrieve secure data that was stored on the cloud. The attack on both systems was successful but PTP system presented the original data that was stored on the cloud while the PNTP system returned a subscriber encrypted data to the hacker. It also took the hacker more time to access files stored on the PNTP system.

4.1. Home Page Interface

Figure 5 illustrates the home page interface, showing the interface of both systems (PTP and PNTP). The home page interface provides subscribers with functionalities like uploading data to the server, encrypting the data, downloading data from the cloud and decrypting the data to get access to the original data. It also shows an evaluation based on time it takes for brute force to access data. The encrypt-and-upload files to server allows the subscriber to make selection of the data to be sent to the CSP and the server encrypt button encrypts the data and saves the encrypted format in the cloud. Access to the data stored on the cloud is possible only when subscriber has the correct encryption and decryption key. On accessing server encrypted data, the subscriber further decrypts data making use of their client decryption key. The second part of the homepage shows the evaluation of both systems making use of a brute force attack. The evaluation allows for the number of simulations to be picked and reports the average success time for each simulation assuming that a hacker has access to the decryption key of the data.

4.2. Interfaces for Dataset File Directory

The program has customized file extensions for encrypted file by the server (.saes) and the client (.caes). The system shared directory for all files encrypted by both the provider and the client as shown in Fig. 6. The caes.saes extension implies that a file was first encrypted by the subscriber before uploaded and was encrypted again by the server of the PNTP system. The encryption and upload to server in Fig. 7 allows subscribers to upload and store encrypted data on the PNTP cloud. A successful encryption notification is returned when encryption is successful otherwise a failure notification is returned. To retrieve data, the subscriber must provide credentials to prove authenticity. At the point of download, the subscriber will have access to the .caes data. A second decryption is done on the subscriber side to retrieve the original highly sensitive data as shown in Fig. 8. Provided that the hackers have access to the correct decryption key, Fig. 9 displays the average time it takes for a hacker to access the files stored on the cloud for both privacy with trusted provider and privacy with non-trusted provider.

5. RESULTS AND DISCUSSION

The system implementation was tested with files that comprises of audio, video, images and text from Stanford University multimedia dataset.

The system implementation revealed the following: client file with size 29666 bytes was successfully encrypted before uploading to NTP system, using 0.608 seconds as shown in Fig. 7. Client file with size 5145440 bytes was successfully decrypted from PNTP system using 1.054 seconds as shown in Fig. 8. In Fig. 10, NTP encrypted client files (caes) with size 48032 bytes using 0.61 seconds. NTP attempted to access client encrypted
Figure 6: Interface for File directory.

Figure 7: Interface for Client to select and upload encrypted files to non-trusted provider.
Figure 8: Interface showing successful decryption of file from non-trusted provider.

Figure 9: Interface for sample server hacking.
file (.caes) with size 48048 bytes using decryption time 0.608 seconds but encountered difficulty accessing the original data because of the client encryption performed on the data.

The system was also evaluated with thirty files from Stanford University multimedia dataset. The proposed PNTP system was evaluated and benchmarked with the PTP system using encryption time, decryption time and efficiency (brute force hacking) as parameters. We assume file sizes ranging from 35 bytes to 200,015,204 bytes and simulation for 10 iterations. The differences in the encryption and decryption time for PTP and PNTP systems are negligible for the same file size as showed in Table 1.

This revealed that, there is no opposition to the use of the proposed PNTP system. The brute force hacking as showed in Table 2 took longer time (almost double) to access data stored on the PNTP system. The results indicated that, it will take a longer time to hack subscriber files if they are first encrypted by the subscriber in the PNTP system before uploaded to a trusted module. Also, the files the hackers get after successfully hacking the PNTP cloud is an encrypted file while that of the PTP cloud is the original file saved on the cloud.

6. CONCLUSION

Cloud computing is embraced by businesses and organizations because of easy access to resources and reduction in the cost of services. While it is important for subscribers to be able to access their information on time and unmodified, it is also important that unauthorized persons do not have access to their information.

Privacy with non-trusted provider (PNTP) solution was proposed to address privacy issues on the cloud platform. The system categorizes subscribers’ sensitivity information into no privacy (NP), privacy with trusted provider (PTP) and privacy with non-trusted provider (PNTP). Privacy with non-trusted provider system was implemented for highly confidential information that has to be kept away from the cloud service providers. The subscriber will have to upload an already encrypted file to the cloud and the file is encrypted again by the service providers. This method is efficient because access to the original file may be difficult. Any decryption attempts by hackers will encounter the encrypted file by the subscriber due to double encryption. The hackers will have to access the subscriber’s encryption key before they can access the original information which is difficult.

Future works include considering other encryption and decryption techniques and compare the
results with AES. The issues of bandwidth, memory and transmission channel consumption can also be considered.

References


Diffie-Hellman key exchange and AES encryption algo-
rithm to enhance data security in cloud computing,” in
IEEE International Conference on Communication Sys-

solution based on fully homomorphic encryption,” in
IEEE 16th International Conference on Advanced Com-

incentive and rewarding scheme for crowd in social me-
dia,” Elsevier Journal of Information Science, vol. 470,

computing system by using encryption and access con-
trol model,” Journal of Information Processing System,