Chapter 3

Veracity Aware Defense Algorithm (VADA) for Cloud computing environment

As discussed in section 1.1, cloud primarily refers to the saving of user’s data to an off-site storage that is managed and controlled by a third party CSP. This means instead of storing data on user computer’s hard disk or other storage devices, the client saves it to a remote database, where the internet provides the connection between the local computer and the remote database.

Computers in the cloud are configured to work simultaneously and the various applications use the collective computing power every bit by using the concept of virtualization. In this model customers plug into the cloud to access all required resources which are priced and provided on-demand. Fundamentally, IT resources are rented and shared like office space, apartments or storage places are utilized by renters, among multiple tenants. Delivered over an internet connection, the cloud eliminates the requirement of company’s own data center. Cloud computing services such as Amazon EC2 and Google App Engine are built to take advantage of already existing huge infrastructure of their respective company.

Additionally to the advantages of cloud computing, there are some shortcomings too. Security in the cloud is one of the main shortcomings in cloud computing because data owner stores his sensitive data to remote hosts and users also access the required information from
remote cloud servers, which is not moderated and managed by data owners himself. The client can ensure the data security using the concept of firewalls, virtual private networks and by other security policies within its own periphery or circumference. But now data is stored outside client premises in the cloud, which raise the issue of data protection. The concept of cloud requires resource sharing such as storage with other cloud vendors; therefore, business critical or other important data of the client is not just available to right cloud vendor but may to third party cloud vendors. So to eliminate the above stated shortcoming, this chapter offers a novel algorithm for cloud environment which includes the three different security systems to accomplish the aim of maximizing the data owners control on data for bringing off the privacy mechanisms and storing, processing and accessing it.

This chapter offers a new algorithm VADA (Veracity Aware Defense Algorithm) for cloud environment which proposes three different security systems to accomplish the aim of maximizing the data owners control on the cloud data. It has been observed that the algorithm is found effective in storing, processing and accessing of different categories of user’s data. The functioning and efficiency of the proposed algorithm are found extremely safe and privacy aware for all types of data on all cloud platforms.

The rest of the chapter has been divided into the following nine sections and presents the approach, developed by the author of this thesis to enhance security of data in existing cloud data storage systems.

Section 3.1 presents the proposed data security algorithm and assumptions taken in advance. Firstly, there are three parties, affected during the process of data storing and accessing, these are- data owner, cloud service provider and cloud user. Secondly, it is taken for granted that each party is preloaded with public keys of others so that there is no need of any public distribution system for the distribution of public keys to each other’s. The proposed algorithm is founded upon these two major assumptions in the cloud.

Section 3.2 presents a simple and effective proposed algorithm VADA for Cloud Environment. The proposed algorithm works in two phases and takes into account the complete control of the user on the data. In the first phase, the cloud data is categorized on the basis its sensitivity and in the second stage the data are transmitted to the CSP according to the sensitivity category of the data, to be stored in the cloud. So before storing and processing the information in the storage pool of CSP, the data owner classifies the data into three classes according to their sensitivity to speed up the operation of data storing and
recovery. It also includes descriptions, detailed Three- Tier Privacy Aware Cloud Computing Model for three different security strategies for each privacy categorized data.

Section 3.3 shows the proposed algorithms for the foremost category of the data, i.e. NP (No Privacy). The section identifies the complete algorithm in detail and all the steps followed. The data are of very little sensitivity and mainly public data, so that plaintext it is transmitted to the CSP through SSL.

Section 3.4 shows the proposed algorithms for the second category of the data, i.e. PTP (Privacy with Trusted Provider). The section identifies the complete algorithm in detail and all the steps followed. The data are of moderate sensitivity and mainly hybrid data, so that plaintext it is transmitted to the trusted CSP through SSL and CSP is responsible for the encryption for the sent data.

Section 3.5 shows the proposed algorithms for the third category of the data, i.e. PNTP (Privacy with Non Trusted Provider). The section identifies the complete algorithm in detail and all the steps followed. The data are of very high sensitivity and mainly private data, so that plaintext is firstly encrypted at user end, then cipher text 1 is transmitted to the CSP through SSL.

Section 3.6 presents the advantages of the proposed algorithms over the traditional solution studied in literature survey.

Section 3.7 presents the shortcomings of the proposed algorithms against the latest attacks identified in cloud data protection. So that the new best approach can be keyed out to protect the cloud data. Two of the identified attacks are likewise identified in detail, e.g. Trivial mathematical attack, Known Plaintext Attack.

Section 3.8 shows the most proposed solutions of the proposed algorithms against the latest attacks identified in cloud data protection.

Section 3.9 summarizes the chapter with the detailed categorization of the data and their individual policy adopted for each category. The advantages, shortcomings and their results are also suggested in the chapter.
3.1 Data security algorithm and assumptions

In VADA algorithm, we assume that the four parties are involved during the communication for data storage and accessing Data owner, cloud service provider, user and trusted module. We also take for granted that each party is preloaded with public keys of others so that there is no need of any PKI for distribution of public keys of each other’s. For large storage and computation capacity, we assume the CSP as a heterogeneous assemblage of several service providers like Google, Amazon and Microsoft.

3.2 Proposed Solution

This section describes the proposed algorithm VADA (VERACITY AWARE DEFENCE ALGORITHM) for Cloud Environment. In which before storing and processing the information in the storage pool of CSP, the data owner classified it into three classes according to their sensibility

(i) No Privacy (NP)
In this category the data is not sensitive and there is no need of any form of encryption. But for network security the data can be sent via SSL.

(ii) PTP (Privacy with Trusted Provider)
In this category data is moderately sensitive and the cloud provider is fully trusted by the data owner. Data owner provides the raw data to trusted provider where the cloud provider itself is responsible for encrypting the data for maintaining its confidentiality and integrity.

(iii) PNTP (Privacy with Non Trusted Provider)
In this category the data are highly sensitive that also needs to be concealed from a cloud provider. This sort of data is encrypted on the data owner side and then stored at the cloud service provider.

The primary focus of VADA is to maximize the data owner’s control in bringing off all views of privacy mechanisms required to preserve the security of sensitive information.

To achieve above defined objective VADA further includes three different protection strategies for each privacy categorized data that has different privacy aspects according to the need of sensitive data. Of this Three- Tier Privacy Aware Cloud Computing Model is proposed for data security shown in figure 3.1 and
The Design framework for implementation of three-tier privacy aware cloud computing model is shown in figure 3.2.
The functional flow diagrams for the NP Privacy category is presented in figure 3.3 and their pseudo codes are written thereafter. On that point is, no encryption and decoding of data during the storage and accessing of No Privacy (NP) category data.

Figure 3.3 - Functional flow diagram of security scheme of NP category

The user just used double encryption during the request made for accessing the data from CSP. The sole prerequisite is that each party must be authenticated before starting their communication.

Whereas in Privacy with trusted provider (PTP) security scheme the CSP is responsible to establish the security of data. CSP used the XOR operation based encryption and decryption technique which automatically gets the key without any complexity. This scheme also has a unique feature of intrusion detection, which prevents from various malicious activities.

3.3 Proposed algorithm for NP Category Data

Step 1: Data owner checks for data privacy category

(a) $\delta pc = NP$

// Here $\delta pc$ represent data privacy category

Step 2: Storage of data in storage pool of CSP

(a) Data owner sends the data to CSP via SSL for network security where data are $\delta DO (NP, data id, data)$

// $\delta DO$ represent data send by the data owner

(b) CSP store data ($\delta DO$) in its storage pool

// Secured data access between user and CSP for NP data category
Step 3: User login to CSP and send request for data access, with user unique ID, access control rights. User encrypts the request with the his private key and public Key of CSP.

Step 4: CSP checks for data category and send data to the user:

(a) Cloud provider receives and decrypt the request with the public key of the user and his own private Key

(b) Check to which category requested data belong

\[ \delta_{pc} = NP \]

//Here data Privacy category is equal to No Privacy (NP)

(c) CSP sends the data to user via SSL for network security where the data is:

\[ \delta_{DO} \]

Step 5: End of algorithm

3.4 Proposed algorithm for PTP Category Data

Step 1: Data owner checks for data privacy category

(a) \( \delta_{pc} = PTP \)

// Here \( \delta_{pc} \) represent data privacy category

Step 2: Storage of data in storage pool of CSP

(a) Data owner sends the data to CSP via SSL for network security where data are \( \delta_{DO} \) (PTP, data id, data)

// \( \delta_{DO} \)represent data send by the data owner

(b) CSP firstly encrypts the data received by using the XOR encryption technique.

(c) CSP store the encrypted data in its storage pool

// Secured data access between user and CSP for PTP data category

Step 3: User login to CSP and send request for data access, with user unique ID, access control rights. User encrypts the request with the his private key and public Key of CSP.

Step 4: CSP checks for data category and send data to the user:
(a) Cloud provider receives and decrypt the request with the public key of the user and his own private Key

(b) Check to which category requested data belong
\[ \delta_{pc} = PTP \]
//Here data Privacy category is equal to Privacy with Trusted Provider (PTP)

(c) CSP decrypts the requested data using the XOR decryption process and sends the data to user via SSL for network security where the data is: \( \delta DO \)

Step 5: End of algorithm

### 3.5 Proposed algorithm for PNTP Category Data

Step 1: Data owner checks for data privacy category

(a) \( \delta_{pc} = PNTP \)
// Here \( \delta_{pc} \) represent data privacy category

Step 2: Storage of data in storage pool of CSP

(a) Data owner first generates a dynamic key from a Trusted Module.

(b) Data owner encrypts the data with any symmetric encryption algorithm using generated key.

(c) Data owner store the generated public key in User/Consumer storage space.

(d) Data owner passes the encrypted data to CSP via SSL where data are \( \delta EDO \) (PNTP, data id, encrypt_data)

// \( \delta EDO \) represent encrypted data send by the data owner

(e) CSP store received encrypted data (\( \delta EDO \)) in its storage pool

// Secured data access between user and CSP for PNTP data category

Step 3: User login to CSP and send request for data access, with user unique ID, access control rights. User encrypts the request with the his private key and public Key of CSP.

Step 4: CSP checks for data category and send data to the user:
(a) Cloud provider receives and decrypt the request with the public key of the user and his own private Key

(b) Check to which category requested data belong
\[ \delta_{pc} = \text{PNTP} \]
//Here data Privacy category is equal to Privacy with Non Trusted Provider (PNTP)

(c) CSP sends the encrypted data to user via SSL for network security where the data is: \( \delta_{E} \text{DO} \)

(d) User decrypts the retrieved data by using the same symmetric decryption process using the same public key stored in user storage space and generate the original data i.e. \( \delta_{DO} \)

Step 5: End of algorithm

### 3.6 Advantages of the algorithm

In the proposed algorithm the XOR operation based cryptography technique generate the key for data packet automatically without any complexity. Its computational cost is nominal as compared to public key encryption when the key size is increased. The XOR operation based cryptography techniques used to behave as intrusion detection and also prevents from various types of intrusions. This also prevents man-in-the-middle attack because it observes the modification of data and even of a bit. In PTP and PNTP, the cloud service provider stores the data in encrypted form which prevent it from inside channel attack. This proposed algorithm ensures required CIA properties of all types of data for all phases.

### 3.7 Shortcomings of the algorithm

The XOR cryptographic operation is relatively light and vulnerable to many attacks as well. Simple mathematical techniques are found those can crack the XOR encryption in a few minutes. To explore the weaknesses of XOR operation, here is a standard algorithms trivial mathematical attack and Known plaintext attack
3.7.1 Trivial mathematical attack

Step 1: Determining of the length of the XOR key: This is done by XORing the encrypted data with itself shifted various numbers of places, and examining how many bytes are same.

Step 2: If the bytes that are equal or greater than a certain percentage, then you have shifted a multiple of the key length. By getting the most modest amount of shifting those results in a heavy measure of equal bytes, you will discover the key length.

Step 3: Shift the cipher text by key length and XOR against itself. This takes out the key and leaves you the plaintext. This is enough plaintext to determine the message contents.

3.7.2 Known Plain Text Attack

This exposure is mainly put forward when one same key is used to code the many files or text. As shortly as we will get to know about one plain text the rest plaintexts can be well obtained. This attacks work in following ways

1. cryptotext = plaintext XOR key
2. key = cryptotext XOR plaintext

Because second statement is mainly equal to key = plain text XOR key XOR plain text. Which will return the key used and if same key has been used with many files then those files also will become vulnerable to the onslaught. This approach is effective enough by having knowledge of some component of the header of plain text file also not full plain text.

3.8 Solutions to the XOR encryption shortcomings

XOR encryption can be reasonably strong if the following conditions are satisfied

(i) The plain text and password should be of about the same length. It increases the key length size, thus it will be harder now to give a trivial mathematical attack.

(ii) Same password should not be used for encrypting more than one message or files. So that, the known plain text attack can be forestalled.

(iii) The XOR key must have one time pad to pull in its length equal to the message and the pad must as random as possible.

(iv) By utilizing the advanced encryption, decryption techniques as employed in next solution for safe cloud storage.
3.9 Summary

In this chapter a new algorithm Veracity Aware Defense Algorithm (VADA) for cloud environment has proposed which included three different security schemes to achieve the objective of maximizing the data owners control for storing, processing and accessing of different categories of private cloud data.

The three parties are affected during the communication for data store and accessing: Data owner, cloud service provider, and user. In traditional cloud data storage policy, sensitive or non sensitive data were only posted to the cloud service provider in plain text by having full trust on the CSP. All cases of raw data were sent to the CSP without much thinking about its secrecy and integrity. The cloud user expected that all the secrecy and integrity must be guaranteed by the CSP, so it is the CSP duty to guarantee it. Only due to some national and external threats identified in the last few years, the user concerned has increased about their data security in the cloud.

The principal focus of proposed algorithm VADA is to maximize the data owner’s control on cloud data. To achieve above defined objective Three- Tier Privacy Aware Cloud Computing Model has proposed for Data Security and the Design framework for implementation of three- tier (No Privacy (NP), Privacy with Trusted Provider (PTP), and Privacy with non trusted provider (PNTP)) privacy aware cloud computing model are offered. So that the most sensitive data will be foreclosed from all national and external approaches.

Thus, in this algorithm the data owner first classifies the information into three categories according to its sensitivity:

(i) NP category
(ii) PTP category
(iii) PNTP Category

In first data category NP (No Privacy), the data is not sensitive (mainly public data) and there is no need of any encryption and decryption. But from network security the data can be sent via SSL. In second data category PTP (Privacy with Trusted Provider), the cloud provider is fully trusted by the data owner. Data owner provides the raw data to trusted provider where the cloud provider itself is responsible for encrypting the data for maintaining its confidentiality and integrity. In third data category PNTP (Privacy with Non Trusted Provider), the data are highly sensitive that also needs to be concealed from a cloud provider.
This sort of data is encrypted on the data owner side too, and then stored at the cloud service provider.

The suggested algorithm has found secure and the proposed concept is further extended in the following chapter to counter the attacks identified on this algorithm.

In next Chapter 4, the cloud storage is assured by using advanced symmetric and asymmetric key cryptographic techniques. We have used 3-DES (well known Symmetric Cryptography technique), MD5 (well known Hashing technique), RSA (well known Asymmetric Cryptography technique) in the succeeding chapter.