Detection and Preventions of Data Leakage in Cloud Computing Environment

Prashant Ingole, Dinesh Jejurkar, Vishal Pawar, K. D. Tamhane
Department Of Information Technology, Pravara Rural Engineering Collage, Loni, Maharashtra.

Abstract – In the recent years internet technologies has become the backbone of any business organization. These organizations utilize this facility to amend their efficiency by transferring data from one location to another. But, there are number of threats in transferring critical organizational data as any culprit employee may public this data. This quandary is kenned as data leakage quandary. In the proposed work, we are suggesting a model for data leakage quandary. In this model, our aim is to identify the culprit who has leaked the critical organizational data.

Index Terms – Bell-LaPadula model (BLP); Hash Function; AES; Watermark; Message chaining.

1. INTRODUCTION

In the current business scenario, data leakage is a sizably voluminous challenge as critical organizational data should be for fended from unauthorized access. Data leakage may be defined as the fortuitous or intentional distribution of private organizational data to the unauthorized entities. It is consequential to for fend the critical data from being misused by any unauthorized use. Critical data include perspicacious copy right information, patent information, functional information etc. In many organizations, this critical organizational data have been shared to many stakeholder outside the organizational premises. Ergo, it is arduous to identify the culprit, who has leaked the data[1].

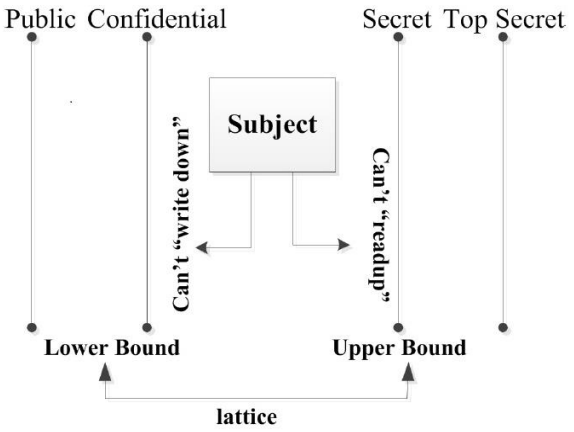

Fig. 1: In the Bell-LaPadula model, each subject S has lattice of rights.

In the proposed work, our goal is to identify the censurable utilizer when the organizational data have been leaked by some agent. In the proposed work, Bell-La Padula security model has been utilized which provide the analysis and design of secure computer systems. This model is called data confidentiality model. Bell-LaPadula model mainly fixates on data confidentiality issues and provides controlled access to relegated information.

In contrast to the Biba-Integrity model which describes rule for the bulwark of data integrity. In this formal model, the entities in an information system are divided into subjects and objects. The notion of a ”secure state” is defined, and it is proven that each state transition preserves security by moving from one secure state to other secure state, thereby inductively proving that the system gratifies the security objectives of the model.

The Bell-LaPadula model is built on the concept of a state machine with a set of allowable states in a computer system. A system state is defined to be secure if the only sanctioned access modes of subjects to objects are in accordance with a security policy. To determine whether a categorical access mode is sanctioned, the clearance level of a subject S is compared to the relegation level of the object O to determine if the subject is sanctioned for the concrete access mode. The clearance/relegation scheme is expressed in terms of a lattice[2] as shown in Figure-1. AES algorithm and RSA algorithm shows good performance among different symmetric and asymmetric encryption technique[3] predicated on different performance factors such as key value, computational speed and tenability. Sundry experimental factors were additionally analyzed predicated on text files used and experimental results proves that DES algorithm consumes least encryption time than AES but in terms of recollection utilization AES uses least time than DES algorithm. In RSA encryption time is more and additionally recollection utilization is very high[2][6]. These techniques are subsidiary for authentic-time encryption. In other model, it has been shown an incipient comparative study between encrypting techniques predicated on nine factors like key length, cipher type, block size, developed, cryptanalysis resistance, security, possibility key, possible ACSII printable character keys, time required to check all possible keys. Predicated on these factors AES is better than DES and RSA. It additionally been discussed that DES is secret key predicated algorithm suffers from key distribution and key acquiescent quandaries but RSA consumes sizably voluminous duration to perform encryption

and decryption operation. It have been additionally observed that decryption of DES algorithm is better than other algorithms in terms of throughput and power consumption[7].

In the recent years, lots of changes transpires in the field of watermarking systems. Digital images are more popular than analog due to facile duplication and transmission on different types of networks. Watermarking is utilized where authentication or ownership is needed[5]. Watermarking is more efficient implement in ownership claiming and fingerprinting of digital data[5][3]. Watermark can be acclimated to transmit secure message from one place.

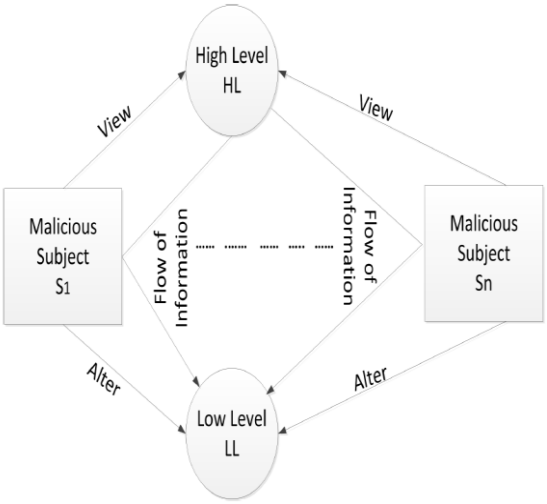

Fig. 2: Information Flow

The computational cost and time involution is the quantification quandaries with robust cryptographic algorithms. These techniques utilize the concept of message authentication. These mentioned techniques ascertain that any transmutation in message can be facilely traced out (active attack) but it fails in case passive attack. Consequently, one single technique is required for both message confidentiality and authentication.

The main concern of the proposed work is to for fend the secret information being transmitted. This paper is structured as follows: In Section II proposed model is discussed. Section III contains Applications and Efficiency Quantification of the proposed model. Results Quantifications and conclusion have been discussed in Section IV and Section V respectively.

2. PROPOSED MODELLING

In this proposed model we are providing the solution for critical Data Leakage quandary. The proposed model has been described in the following sections-

A. Secured Environment Infrastructure:-

We are utilizing the concept of Bell-LaPadula Model for providing secured infrastructure,

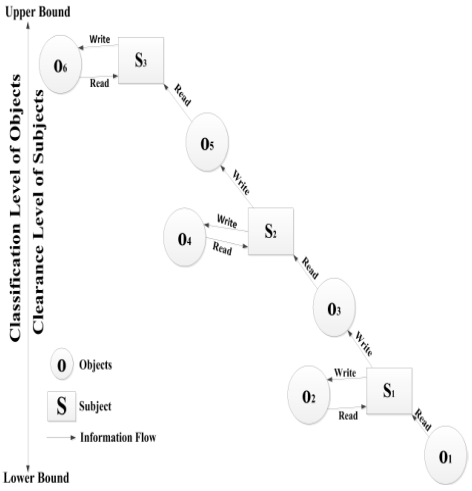

Fig. 3: Read and Write access provided by the Bell-LaPadula model

It is a state-machine model and used to apply access control in different environment such as- Military security - Army, Air-force, Navy, NATO, NASA etc. Commercial security- Marketing Sales, Research and development, Human Resource department etc. In Bell-LaPadula model, information flow will be between the high calibers to low caliber it is shown in Figure-2. We define state, if the system as a secured environment, and it follows some rule defined, as the sanctioned access mode the any subject S, with the any object O is sanctioned, with reverence to defined security policy. To find whether any concrete access mode will be sanctioned, the clearance of a subject S is compared to the relegation of the object O. i.e. S =(S1, S2, S3,..Sn), O =(O1, O2, O3,..On) both S and O are coalesce and engendering up the security level used to determine if the subject S is sanctioned for the concrete access mode[4].

A. Engineering Watermark:-

In this model server will integrate an image logo to all the stored documents and this image logo represents the organization For each of the three components of color image as RED, GREEN and BLUE ranges from 0 to (28□1). Each character, has their ASCII values ranges from 0 to (28 □1). So, any text can be Inserted into the document by superseding the intensity value of pixel location, with the ASCII value of character, which is needed to be obnubilate and transmit with the document So, first applying any cryptographic algorithm to text and then embedding the resulting ASCII to document. This process ascertains the indispensable security. The key conception abaft the implementation of this technique is to embed secret message into the document with a computationally secured and time efficacious manner. In lieu of utilizing a high weighted cryptographic algorithm like RSA,

an efficacious light weighted algorithm like AES can be utilized with authentication scheme like SHA- 512[5][3].

The main focus of the proposed model is that only the registered utilizer will be able to access the critical document otherwise non-registered utilizer has to first register itself with the server it is shown in Figure-4. This Watermarking technique discuss how and where to place the authentication code in the critical document D. The server securely will maintain a server directory table for each registered client's id which is shown in TABLE-1. The input to the algorithm is pristine document D, secret message (IDC) and 128 bit key (K) utilized in AES- 128 encryption scheme which engender the cipher text C as an output. In second phase of watermark embedding includes another input IV , (initial value), which is 512 bits long, and used for engendering 512 bits long message authentication code M, it embedded into the document D. The process of watermark embedding in document is describe in following phases

1) Phase I: Calculation of all parameters:

Calculate the cipher text, C, by utilizing secret message (IDC), Encryption Key K and AES-128 encryption Algorithm. It will be implemented utilizing block cipher techniques[6][2].

Calculate message authentication code, M, utilizing IDC, initial vector (IV) and SHA-512 scheme. Notice that, M is engendered by utilizing secret message (IDC), not by utilizing the cipher text C. This will discombobulate the intruder, and will provide the extra level of security [16].

Calculate situating pixels in the document D as:

◦ Row situating pixel, m = I(1; 1) + 2

◦ Column situating pixel, n = I(1; 2) + 2

2) Phase II: Placement of cipher C and authentication code M into image:

Supersede the pixel value starting from (m, n) in the pristine document, with the value of cipher text C. Each block in cipher text will transmute precisely 16 pixel bits in the pristine document D.

Supersede the last 64 pixel bits, with the authentication code M calculated, in inversion order. This will perplex to intruders and provide the extra level of security. Determinately the watermarked document WMD will be engendered as the output for this process as shown in Fig

B. Sending WMD To Client-

In this phase the engendered Watermarked document WMD will be send to the requested client along withserver's public key certificates(PKCserver), which verify integrity of the genuine source of document. This Server will utilize the nonce (Cnonce) in order to forefend the man in middle attack. The send document will be encrypted with the public key of the client (PUC) and contain the hashS which is engendered by the server. The process of sending WMD to client is shown in Figure.

Client will receive the send document and open with the avail of his private key PRC and verify the document, by engendering the hash of received WMD document. If the received hashS is identically tantamount to the engendered hashC (ie. hashS = hashC) then the document is not altered in between and document integrity is maintained.

Analyze the secret. The proposed technique can be utilized with the variants of image such as Binary images, Gray scale images and Color images etc. so it can be verbally expressed that this technique is best suited to transmit symmetric keys in secure manner

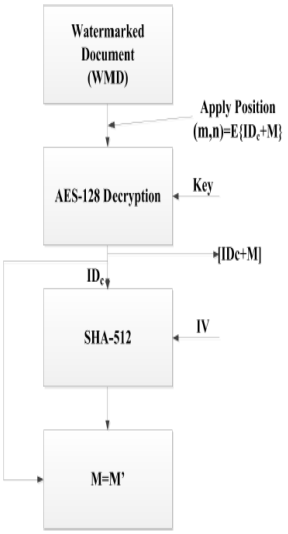

Fig.: 4 The Process client id detection form WMD

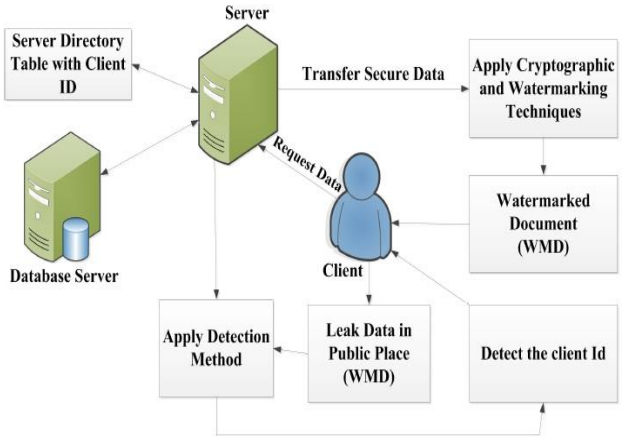

Fig. 5: Overall Working of the Proposed Model

The technique is very economical, because it utilizes light cryptographic algorithm AES-128 with SHA-512 to provide double security with half computational time[6]. If RSA is utilized then it leads to two major quandaries[7]. First its key length is very high, i.e., of 1024 bits, Second it will be more arduous to calculate exponential computations than simple computations needed in AES which utilizes the key of length only 128 bits long. The technique proposed in this paper utilizes the commixing concept to engender message authentication code, and put it in inversion order to perplex the intruders.

3. APPLICATION AND EFFICIENCY QUANTIFICATION

The proposed method is efficient to utilize with any size of documents. Here, we have utilized only client IDC with the pristine document. Ergo, in this case if the obnubilated message is short, the vicissitudes made to the pristine document is withal short, resulting in less vicissitude in pristine document, and hence intruder will not be able to analyze the secret. The proposed technique can be utilized with the variants of image such as Binary images, Gray scale images and Color images etc. so it can be verbally expressed that this technique is best suited to transmit symmetric keys in secure manner.

The technique is very economical, because it utilizes light cryptographic algorithm AES-128 with SHA-512 to provide double security with half computational time[6]. If RSA is utilized then it leads to two major quandaries[7]. First its key length is very high, i.e., of 1024 bits, Second it will be more arduous to calculate exponential computations than simple computations needed in AES which utilizes the key of length only 128 bits long. The technique proposed in this paper utilizes the commixing concept to engender message authentication code, and put it in inversion order to perplex the intruders.

4. RESULT MEASURMENT

In this model we have culled the AES model because it is more expeditious in the process of encryption and decryption. To crack the 128-bit AES key utilizing a well-kenned brute force attack it would take 1 billion years. AES is the successor of DES as standard symmetric encryption algorithm for US federal organizations. AES accepts keys of 128, 192 or 256 bits (128 bits is already very unbreakable), uses 128-bit blocks (so no issue there), and is efficient in both software and hardware. It was culled through an open competition involving hundreds of cryptographers during several years. If we compare these algorithms in terms of encryption and decryption we find that time taken in AES encryption of message over different size of the message is less than DES and RSA.

It can be verified in Table-I and Table-II by comparing the encryption and decryption time of AES with DES and RSA Algorithm[7][8].In the Table-II it shows the decryption time for different size of the messages AES takes less time over DES and RSA Algorithm. In Table-III where we analyze the different factors which will shows the characteristics of the DES, AES and RSAAlgorithms[8].

SR.NO	DES	AES	RSA	Data Size
1	3.0	1.6	7.3	153KB
2	3.2	1.7	10.0	118KB
3	2.0	1.7	8.5	196KB
4	4.0	2.0	8.2	868KB
5	3.0	1.8	7.8	312KB

TABLE I: Comparison of various packet sizes for DES, AES & RSA algorithm (Encryption Time)

SR.NO	DES	AES	RSA	Data Size
1	1.0	1.1	4.9	153KB
2	1.2	1.2	5.0	118KB
3	1.4	1.24	5.9	196KB
4	1.8	1.2	5.1	868KB
5	1.6	1.3	5.1	312KB

TABLE II: Comparison of various packet sizes for DES, AES & RSA algorithm (Decryption Time)

Factors analyzed	DES	AES	RSA
Development Years	1997	2000	1978
Key-Length (Bits)	56	128,192, 256	<1024
Nature of Algorithms	Symmetric	Symmetric	Asymmetric
Encryption/Decryption (Speed)	Low	High	Medium
Nature of Security Attacks	Inadequate	Highly Secured	Highly Secured

TABLE III: Analysis of various factors

5. CONCLUSIONS AND FUTURE SCOPES

The proposed technique will provide better security against data leakage quandary.We can detect the data leaker in authentic time by utilizing this method. It additionally bulwark variants of active and passive attacks. The proposed technique is computationally cost efficacious in terms of time and space uses. Ergo, this can be subsidiary in distributed computing

environment to bulwark data from data leakage. The proposed technique is predicated on symmetric algorithm, consequently it is infeasible to elongate this model for web environment where multiple number of users frequently accessing the data object. We can additionally implement this technique for asymmetric cryptography.

REFERENCES

[1] Rupesh Mishra and DK Chitre. Data leakage and detection of guilty agent. International Journal of Scientific & Engineering Research.
[2] David Elliott Bell. Bell–la padula model. Encyclopedia of Cryptography and Security, pages 74–79, 2011.
[3] JJK RUANAIDH and T PUN. Rotation, scale and translation invariant spread spectrum digital image watermarking. Signal processing, 66(3):303–317, 1998.
[4] Mukesh Singhal and Niranjan G Shivaratri. Advanced concepts in operating systems. McGraw-Hill, Inc., 1994.
[5] Achal Kumar and Vibhav Prakash Singh. Digital watermarking using color image processing using images for transmitting secret information.
[6] E Thambiraja, G Ramesh, and Dr R Umarani. A survey on various most common encryption techniques. International journal of advanced research in computer science and software engineering, 2(7):226–233,2012.
[7] Aman Kumar, Sudesh Jakhar, and Sunil Makkar. Distinction between secret key and public key cryptography with existing glitches. Indian Journal of Educati.