Journal of Network and Computer Applications 35 (2012) 1831–1838

Contents lists available at SciVerse ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

A combined approach to ensure data security in cloud computing

Sandeep K. Sood *

Department of Computer Science and Engineering, GNDU, Regional Campus, Gurdaspur (Punjab) 143521, India

ARTICLE INFO

Article history:
Received 16 January 2012
Received in revised form
25 May 2012
Accepted 3 July 2012
Available online 25 July 2012

Keywords:
Cloud security
Encryption
Message authentication code
Virtualization
Secured socket layer

ABSTRACT

Cloud computing is a forthcoming revolution in information technology (IT) industry because of its performance, accessibility, low cost and many other luxuries. It is an approach to maximize the capacity or step up capabilities vigorously without investing in new infrastructure, nurturing new personnel or licensing new software. It provides gigantic storage for data and faster computing to customers over the internet. It essentially shifts the database and application software to the large data centers, i.e., cloud, where management of data and services may not be completely trustworthy. That is why companies are reluctant to deploy their business in the cloud even cloud computing offers a wide range of luxuries. Security of data in cloud is one of the major issues which acts as an obstacle in the implementation of cloud computing. In this paper, a frame work comprising of different techniques and specialized procedures is proposed that can efficiently protect the data from the beginning to the end, i.e., from the owner to the cloud and then to the user. We commence with the classification of data on the basis of three cryptographic parameters presented by the user, i.e., Confidentiality (*C*), Availability (A) and Integrity (I).The strategy followed to protect the data utilizes various measures such as the SSL (Secure Socket Layer) 128-bit encryption and can also be raised to 256-bit encryption if needed, MAC (Message Authentication Code) is used for integrity check of data, searchable encryption and division of data into three sections in cloud for storage. The division of data into three sections renders supplementary protection and simple access to the data. The user who wishes to access the data is required to provide the owner login identity and password, before admittance is given to the encrypted data in Section 1, Section 2, and Section 3.

1. Introduction

Cloud primarily refers to saving of user's data to an offsite storage system that is maintained by a third party. This means instead of storing information on user computer's hard disk or other storage devices, client save it to a remote database where internet provides the connection between user computer and the remote database. Computers in the cloud are configured to work simultaneously and the various applications use the collective computing power as if they are running on a cloud using the concept of virtualization. In this model customers plug into the cloud to access information technology resources which are priced and provided on-demand. Essentially, IT resources are rented and shared among multiple tenants like office space, apartments or storage spaces are used by tenants. Delivered over an internet connection, the cloud eliminates the company's data center or server. Cloud computing services such as Amazon EC2 and Google App Engine are built to take advantage of the already existing infrastructure of their respective company.

The cloud computing model revolves around three functional units or components as listed below:

1. **Cloud service provider:** It is an entity, which manages Cloud Storage Server (CSS), has significant storage space to preserve the clients' data and high computation power.
2. **Client/owner:** It is an entity, which has large data files to be stored in the cloud and relies on the cloud for data maintenance and computation; it can either be individual consumer or organizations.
3. **User:** It is a unit, which is registered with the owner and uses the data of owner stored on the cloud. The user can be an owner itself as well.

Cloud can ensure the user's data security using the concept of firewalls, virtual private networks and by implementing other security policies with in its own periphery or perimeter. Since the concept of cloud requires resource polling with other cloud owner's, hence, business critical or other important data of client is not only available to cloud but also to third party cloud (Julisch and Hall, 2010). Security is therefore a major element in any cloud computing infrastructure, because it is essential to ensure that only authorized access is permitted and secure behavior is expected (Overby et al.,

* Tel.: +91 1894276861.
E-mail address: san1198@gmail.com

1084-8045/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jnca.2012.07.007

2006). The various security concerns and upcoming challenges are addressed in (Daniel and Wilson, 2003; Dikaiakos et al., 2009) and also reviewed in terms of standards such as PCI-DSS, ITIL, and ISO-27001/27002. There are also architectural security issues which are changing according to various architectural design functioning over cloud computing. Since outsourcing is the main theme of cloud computing, there are two main concerns in this area:

1. External attacker (any unauthorized person) can get to the critical data, as the control is not in the hands of the owner.
2. Cloud service provider himself can breach the owner, as data is to be kept in his premises.

Any kind of security and privacy violation is critical and can produce dire consequences. As soon as cloud privacy issues are further organized and strict regulations and governance for cloud operation are in position, more and more business owners will feel safe to opt for cloud computing.

The proposed model has been structured by bringing together various techniques and utilizing them to perform the task of data security in cloud. This combination of diverse methods operate as a wall stood together against the security challenges, which have been constantly creating the loop holes in the efficient functioning and growth of the cloud. This model is described in such a way that it provides a complete view of processing the data at different levels. The model uses encryption as the main fundamental protection scheme and data sent to cloud is in encrypted form. Encryption is the conversion of data into encrypted form called a cipher text that cannot be easily understood by unauthorized person and can be decrypted by the authorized person having a valid decryption key. Apart from this, the model positively handles the security issues by employing strict authentication parameters, digital signatures, storing encrypted data in cloud according to sensitivity rating, building of index, using of MAC for integrity check and keyword search for data in cloud. Thus all these parameters result into a defined mechanism that encourages the proper functioning of cloud computing. In this computing model, owner sends the encrypted data to cloud where it is stored in different sections depending on the sensitivity rating and then the data can be retrieved by user from the cloud when requested. However, this is achievable only after passing the authentication parameters and then searching the data by the use of keyword obtained from the owner.

This paper is structured as follow: Section 2 summarizes the related work for security of data. In Section 3, a model is proposed which is designed to solve the security issue of cloud computing. Section 4 provides the security analysis of the designed model. Section 5 compares functionality of proposed model with other security models. Section 6 shows the evaluation procedure and Section 7 concludes this paper.

2. Related work

The cloud is a terminology with a long history in telephony, which has in the past decade, been adopted as a metaphor for internet based services, with a common depiction in network diagrams as a cloud outline. The underlying concept dates back to 1960 when John McCarthy opinion that "Computation may someday be organized as a public utility"; indeed it shares characteristics with service bureaus which date back to the 1960s. The term cloud had already come into commercial use in the early 1990s to refer to large Asynchronous Transfer Mode (ATM) networks. By the turn of the 21st century, the term "cloud computing" had started to appear, although major focus at this time was on Software as a Service (SaaS). In 1999, salesforce.com was established by Marc Benioff, Parker Harris. They applied many technologies of consumer web sites like Google and Yahoo! to business applications. They also provided the concept of "On demand" and "SaaS" with their real business and successful customers. IBM extended these concepts in 2001, as detailed in the Autonomic Computing Manifesto, which describes advanced automation techniques such as self-monitoring, self-healing, self-configuring and self-optimizing in the management of complex IT systems with heterogeneous storage, servers, applications, networks, security mechanisms and other system elements that can be virtualized across an enterprise. Amazon.com played a key role in the development of cloud computing by modernizing their data centers. It found that the new cloud architecture resulted in significant internal efficiency improvements and providing access to their systems by way of Amazon Web Services in 2005 on a utility computing basis. 2007 saw increased activity with Google, IBM and a number of universities embarking on a large scale cloud computing research project, around the time the term started gaining popularity in the mainstream press. In August 2008, Gartner Research observed that "organizations are switching from company-owned hardware and software assets to per-use service-based models". The projected shift to cloud computing will result in dramatic growth in IT products in some areas and in significant reductions in other areas. Despite all the hope of gaining maximum advantage from this cloud computing, it seems to have born with security and management concerns, which time to time hinders its growth. For this, lot of research work has been done to secure the data in cloud computing (primary concern) from every perspective, but everything seems to face a new challenge as soon as it is employed.

Juels et al. (2007) described a formal Proof of Retrievability (POR) model for ensuring the remote data integrity. Their scheme combines spot-checking and error-correcting code to ensure both possession and recovery of files on archive service systems. Shacham and Waters (2008) built on this model and constructed a random linear function based Homomorphic Authenticator. This enables unlimited number of queries and requires less communication overhead. Bowers et al. (2008a) proposed an improved framework for POR protocols that generalizes both Juels and Shacham's work. Later in their subsequent work, Bowers et al. (2008b) extended POR model to distributed systems. However, all these schemes are focusing on static data. The effectiveness of their schemes rests primarily on the pre-processing steps that the user conducts before outsourcing the data file. Any change to the contents of data file, even a few bits, must propagate through the error-correcting code, thus, establishing significant computation and communication complexity.

Chor et al. (1995) proposed private information retrieval (PIR) so that clients can access entries in a distributed table without revealing which entries they are interested in. The PIR literature usually aims for very strong information theoretic security bounds, which makes it harder to find practical schemes. PIR schemes often require multiple non-colluding servers, consume large amounts of bandwidth, do not guarantee the confidentiality of the data, do not support private keyword searching and do not support controlled searching or query isolation. The schemes (Cachin et al., 1999; Chor et al., 1998; Gertner et al., 1998; Kusilevitz and Ostrovsky, 1997) are important exceptions which allow removing some but not all these limitations.

Recently, Wang et al. (2009) described a homomorphism distributed verification scheme using Pseudorandom Data to verify the storage correctness of user data in cloud. This scheme achieves the guaranty of data availability, reliability and integrity. However, this scheme was also not providing complete protection to user data in cloud computing, since pseudorandom data would not cover the entire information.

Prasad et al. (2011) and Sood et al. (2011) discussed different security aspects in computing. Prasad et al. (2011) technique provides a new way to authenticate in 3-dimensional approaches. It provides

availability of data by surmounting many existing problem like denial of services and data leakage etc. Additionally, it also provides more flexibility and capability to meet the rising demand of today's complex and diverse network. But in this model, the data stored is not in encrypted form and once the username and password is lost, the data can easily be retrieved by any unauthorized user.

Kamara and Lauter (2010) worked over public cloud infrastructure and proposed a model which is well suited for preserving integrity with the help of cryptographic primitives. This technique is purely based on cryptographic storage services. In proposed procedure, when a user wants to send data to other user, they first generate a master key that encrypts their message. The secret key for decryption is stored on receivers' system for decrypting the same message. They use the concept of index encryption and tokens are generated with the knowledge of secret key. The searching method is not very efficient for encrypted data. They discussed symmetric searchable encryption (SSE) and asymmetric searchable encryption (ASE). These techniques are used for encrypted data searching but increase complexity and make the system cumbersome.

Wang et al. (2010) discussed the drawbacks of using ordinary encryption techniques and suggested that these techniques are not useful over cloud because for this user should have pre knowledge about the encrypted cloud data. Their model is based on symmetric searchable encryption method. They gave design for existing cryptographic primitive and order preserving symmetric encryption (OPSE). Security analysis shows its success rate for one to many mapping and for ranked keyword search. This model did not provide any information about the security attacks, confidentiality and integrity. This model is not well suited for preserving security.

Popa et al. (2010) presents Cloud Proof, a secure storage system for increasing security over cloud. In this model users can detect violations of integrity, confidentiality, write serial ability and freshness. Model use cryptographic tools and engineering efforts to obtain an efficient and scalable system which allow users to detect and prove cloud misbehavior.

Cloud computing is a layered technology and the data in cloud computing has to go through different processing levels, so the security mechanism should be efficient and provided at each step, i.e., from owner to cloud and cloud to user or back to owner. Data should not succumb to the attackers trying to retrieve or tamper with it and not even the cloud provider should be able to harm the data in any possible manner, because cloud service provider cannot be trusted with data of high sensitivity. Hereby we can say that the proposed model has been designed by keeping all these things in mind and surely in comparison to prior works, provides all these required measures to protect data in a very efficient and organized manner.

3. Proposed model

Proposed framework has been structured to provide complete security to the data throughout the entire process of cloud computing, be it in cloud or in transit. Thus, multiple mechanisms and available techniques are applied to shield the critical information from unauthorized parties. The proposed frame work is divided into two phases. First phase deals with process of transmitting and storing data securely into the cloud. Second phase deals with the retrieval of data from cloud and showing the generation of requests for data access, double authentication, verification of digital signature and integrity, thereby providing authorized user with data on passing all security mechanisms.

3.1. Phase 1(storing of data)

This phase deals with mechanisms and methods to store and secure the data from beginning and transmitting it securely to the cloud in encrypted form. It is further divided into sub-sections (Classification, Index Building and encryption, Message Authentication Code (MAC) which provide stepwise details of action on the data).

3.1.1. Classification

As the data in the cloud is intended to be stored, an approach is introduced for storing the data in different sections in the cloud (*public*, *private*, *limited access*) basis of three cryptographic parameters viz: Confidentiality, Availability and Integrity. These values will be listed by the client himself and sensitivity rating (*SR*) will be calculated using the proposed algorithm shown ahead. The value of *C* (confidentiality) is based on the level of privacy needed at each step of data processing, value of *I* (integrity) is based on how much accuracy of data, reliability of information and protection from unauthorized modification is required, and value of *A* (availability) is based on how frequently data is accessed and should available immediately when requested Fig. 1.

Algorithm.

1. Input: Data, protection section, D [] array of n integer size.
   Where D[ ] array consisting of C, I, A, SR, R of n integer size.
2. Output**:** Categorized data for corresponding section.
3. For i=1 to n
   - 3.1 C [i]=Value of Confidentiality.
   - 3.2 I [i]=Value of Integrity.
   - 3.3 A [i] =Value of Availability.
   - 3.4 Calculate *SR* [i]=(C[i]+(1/A[i])*10+I[i])/2 /*security and confidentiality is directly proportional to integrity and availability is inversely proportional to security*/

4. For j=1 to 10
   - For i=1 to n
   - IF SR [i]== 1||2||3 then /* where || represents OR operation.
   - S[i]=3 /* Section 3 allotted to D[i]th data.
   - IF SR[i]== 4||5||6 then
   - S[i] =2 /* Section 2 allotted to D[i]th data.
   - IF SR [i]== 8||9||10 then
   - S[i] =1 /* Section 1 allotted to D[i]th data.

In the algorithm listed above, the primary job of the owner is to categorize the data on the basis of cryptographic parameters viz: *C*, *I* and *A*. Here *D* [ ] represents the data and the user has to give values of *C*, *I* and *A*. After applying the proposed formula as shown above, the value of Sensitivity Rating (*SR*) is calculated. This "*SR*" value is used to allocate the data to one of the three sections in cloud, i.e., S3 [Public], S2 [Private] or S1[Owner's Limited Access] as shown in Fig. 2.

3.1.2. Index building and encryption

After the successful allotment of values to data, the data now needs to gear up for another processing mechanism. As the data on cloud will be stored in encrypted form and searching over encrypted data is a complicated issue, so we need to build up an index, using Index builder shown in Fig. 3, so that while retrieval, we can perform searching over encrypted data. Possible way to build up an index is that, for each word W (keyword) of interest, list the documents that contains W. Building up an Index provides faster retrieval of files. To provide more security against revealing any sort of information to cloud we will encrypt the index also. This index will basically contain a list of keywords, with each keyword contains list of pointers to the documents where keyword appears. The keywords are words of interest that a user may want to search later. Best practice is to build an index of clear documents and then encrypt both document and index and store the encrypted data onto the cloud. The index should be encrypted, by encrypting keywords as well as document pointers in each list in the index. After this we need to encrypt data. Now, to code the data, the model uses encryption. Encryption is the process of turning intelligible information into useless information. With Secure Socket Layer (SSL) encryption, there is also a key that allows only an authorized person to be able to decode the information. This model uses 128-bit SSL encryption to encrypt the data as well as index as shown below.

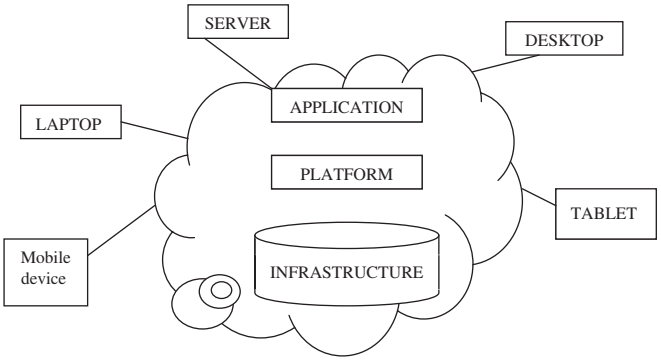

Fig. 1. Concept of cloud.

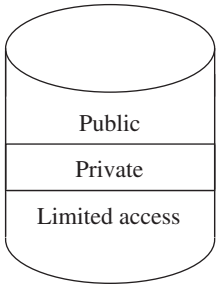

Fig. 2. Data segregated into three sections—S3 [Public], S2 [Private], S1 [Owner].

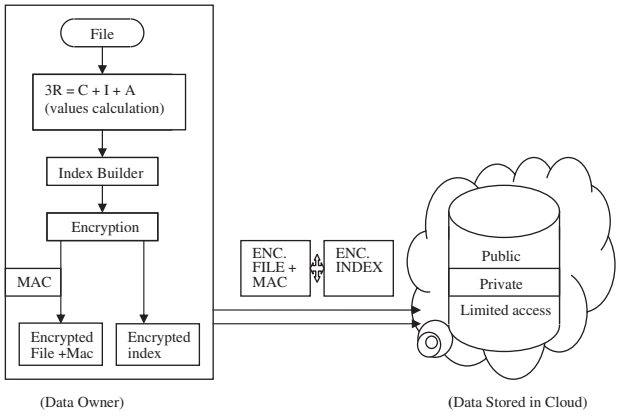

Fig. 3. (Data owner) (Data stored in cloud).

$$128\text{-bit SSL}\ (F,k_2)\rightarrow F'' \tag{1}$$

As shown above we will use any key (k_2) to encrypt the file *F* and results into F'' (encrypted file) and this F'' can be decrypted only by using the same key k_2. Fig. 3 show data transfer from owner to cloud.

If we look at 128-bit encryption using SSL, we will see that there are 2^{88} more bits of key length than previous 40-bit encryption of SSL. Just that change means that there are 2^{88} more combinations. This makes it much harder for hackers to try to crack the code. The value range is beyond the range of the trillions. So, that means 256-bit encryption using SSL is stronger. For the most part, 128-bit encryption is more than sufficient. It is complex enough to make a brute force attack. The processing power needed, among other things, would render most attackers ineffective. However, as technology advances, it is expected that at some point the industry standard will have to shift to 256-bit encryption for SSL. As the size of key increases the cost increases but not significantly, there is cost and key size trade off exists.

3.1.3. Message authentication code

After encryption the data, a message authentication code (MAC) is generated which it transmits along with the encrypted data to cloud. MAC is a small fixed size block of data that is generated based on message/file *F* of variable length using any secret key. It is called cryptographic checksum and is used to check whether data has been tampered throughout the transmission and this check can be made by the user or owner of data on retrieving the file. Fig. 4 shows the individual working of generating MAC of file F'' using key k_1 and providing the encrypted data and its MAC on the other side.

With the generation of message authentication code, the model has data ready to be sent to the cloud for storage as shown in Fig. 4. Now, as the encrypted data on reaching cloud is to be stored in particular sections, the data will be distinguished on the basis of the Sensitivity Rating calculated i.e., $SR \leq 3$ will go into public sections (S3), $3 < SR \leq 6$ will go into private sections (S2) and $6 < SR \leq 10$ will go into owner's limited access sections (S1).

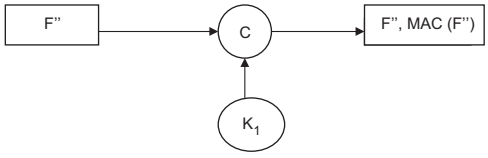

Fig. 4. Generating MAC.

3.2. Phase (retrieval of data)

Now when the data has been stored in cloud in secure manner, the retrieval of data should be supported with equally best possible mechanism and techniques. First the retrieval of data requires the user to register him with the owner/organization by getting a username and a password as shown in Fig. 5. The user will register to get its username and password at organization, which will further forward the username to cloud to let it store the username into its directory.

In this model, when the user requires accessing the data in cloud, he sends a request along with the username to cloud. Cloud check the request and if it is for public section (Section 3), then without authentication access is granted and user after retrieving can decrypt this data by the public key provided in the section only. If the request is for private section (Section 2) and limited access section (Section 1), authentication is necessary and cloud looks for username provided by user into its directory of usernames, provided by the owner. Fig. 6, given below, depicts the transmission of these requests. The sectional database used in this model will provide access under the following guidelines:

1. The user granted access to upper section is not allowed to access lower section, i.e., no read/write is in lower section. e.g.: if user has got access for data in public section, then the data of same owner available in private and limited access section will not be available to the user.
2. The user granted access to lower section is allowed to access upper section. e.g.: if user has got access for data in limited access section, then the data of same owner available in private section will also be available to the user.

As seen in Fig. 7, if the username matches, cloud forwards the username to the owner/organization for authentication, here the important thing is that primary authentication process is kept with the owner only because criticality of the process is such that even the cloud cannot be trusted. Now as the company receives the username from cloud, it has to authenticate the username.

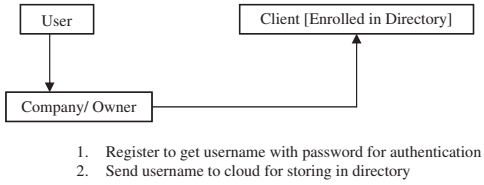

Fig. 5. Registration process.

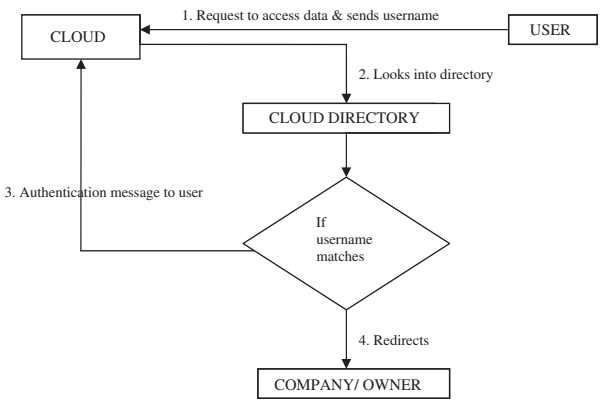

Fig. 6. Access request.

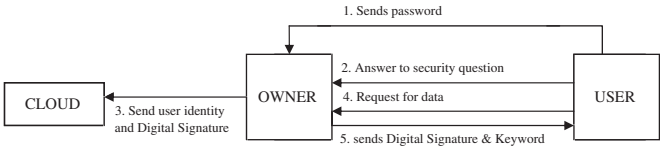

Fig. 7. Authentication process.

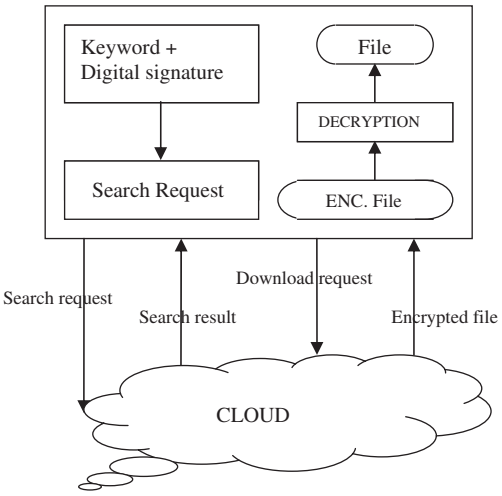

Fig. 8. File retrieval process.

Fig. 7 shows the authentication and request of user to owner for data.

For authentication, the user first sends the password to the owner and on clearing this parameter, user is prompted with a security question from owner and after answering it correctly, user is authenticated. The owner also sends the user identity along with the digital signature to cloud so that cloud will be sure that owner has given the access of data to particular user identity for this session only. Afterwards user sends the request for data to owner which in turn sends the owner's "Digital Signature", keyword of requested data and a master key to decrypt the data provided by cloud. On receiving the digital signature and keyword from the owner, the user forwards the same to the cloud with a search request for data corresponding to the keyword as shown in Fig. 8. The cloud first verifies the digital signature, and if verified, cloud processes the search request using the keyword. Basically searching over encrypted data provides easy retrieval of file and without revealing any critical information to the cloud. As explained earlier we already have stored an encrypted index containing a list of keywords and with each keyword list of pointers to the document where keyword appears. Whenever the cloud gets a keyword to search over encrypted data, it finds a match and then returns the user encrypted list of matching positions from the index. The user can then decrypt the encrypted entries by the decrypting key provided by owners of data and sends cloud download request to retrieve relevant file or document which user was looking for. The cloud replies the user with encrypted file it requested for and then user can decrypt the file by decryption key of file F, already shared by owner with user. One possible advantage for this scheme is that the request could be embedded in other retrievals so that cloud might have uncertainty about the correlation of the search request and the retrieval request for cipher text. Fig. 8 shows the user generating the search request and retrieving encrypted file from the cloud.

Now, as the user has received his data from cloud, the question or doubt arises in the mind regarding its integrity. As this model uses MAC for integrity check, the user can assure him by deriving

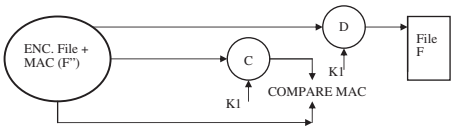

Fig. 9. MAC comparison.

the MAC of encrypted file received using secret key, already shared with user and comparing it with MAC received along with the encrypted file. Here the verification process of data integrity can be performed as shown in Fig. 9.

Fig. 9 demonstrates that when the encrypted file is retrieved by user, Key (k_1) is used to derive its MAC^ and then compared with the MAC received along with the encrypted data. If MAC^=MAC, i.e., MAC of encrypted data which was computed by owner and MAC^ of encrypted data computed by user is equal then it proves that data has not been tampered throughout the whole data traversing. On the other hand, the user can use the key (k_1) to decrypt the encrypted file received. With this step the process of retrieval concludes and we have seen that throughout the whole working, the model has taken all the required precautions and measures to protect the data from possible attacks such as data leakage, unauthorized access and tampering of data etc.

4. Security analysis

The analysis of the proposed model for security of data throughout the whole traversing into this cloud computing paradigm comes up with the following mentioned steps where data can be very vulnerable to threats like data leakage, modification, privacy of users and confidentiality etc. The proposed model is designed to tackle all these security issues very efficiently.

4.1. Unauthorized server

As the data needs to be transmitted over a network to the cloud, there are numerous means through which an attacker can easily get into the internet based network and act as a cloud server to the owner of data, hence resulting into the loss of data. To prevent the loss of data in this situation, SSL certification in this model is used. Certificate Authorities (CAs) issue each certificate, which is a credential for the online world, to only one specific domain or server. The cloud server first sends the identification information to the owner when it connects then sends the owner a copy of its SSL Certificate. The owner verifies the certificate and then sends a message to the server and the server sends back a digitally signed acknowledgement to start an SSL encrypted session, enabling encrypted data transfer between the browser and the server. Moreover, the data and keywords are stored on the cloud in encrypted form.

4.2. Brute force attack

The data while in transmission to cloud over an internet network can be attacked by various unauthorized interceptors. Since SSL offers encryption that prevents interceptors from reading data traversing the cloud. It is not difficult to crack using today's computers which can crunch large number combinations quickly in order to determine every possible key in an effort known as a brute force attack. Thus, in proposed model we are using 128-bit SSL encryption which provides more bits of key length than the previous one SSL (40 bit) and also can be shifted to 256-bit whenever required. 128-bit SSL is complex enough to make a brute force attack mostly useless at this time. The proposed model uses double encryption, one being done by owner and other using SSL. The processing power needed, among other things, would render most attackers ineffective. Hence this approach not only safeguards data where it lives, but also helps assure customers that data is secure while in transit.

4.3. Threat from cloud service provider

The cloud is the place where the data resides after being transmitted by the owner. Suppose the data in cloud is safe from any third party, as the cloud service provider will use strict measures to protect it. The cloud service provider can turn against the owner. As the data is not in the control of owner when in cloud, anything can be possible or cloud service provider can manage any leakage of data even by helping the rival parties. So, the cloud service provider (CSP) cannot be trusted blindly. For this the best possible solution used in proposed model is encryption of data stored in cloud. SSL Certificates as used in the proposed model encrypts private communications over the public Internet. Using public key infrastructure, SSL consists of a public key (which encrypts information) and a private key (which decrypts information), so that only the key owners can read it. 128-bit SSL encryption encrypts the data in such a way that it is nearly impossible for an attacker to decrypt it by a brute force attack.

4.4. Tampering of data

The data is always under the threat of being tampered by any unauthorized interceptor. As all the precautionary measures such as encryption of data, keywords and SSL encryption have been taken in the proposed model to not let anyone tamper the data, but still data needs to be checked after the transmission. For this, MAC (Message authentication code) has been used in proposed model. MAC of encrypted data is generated by the owner before sending it and this MAC is transmitted along with the encrypted data. On the other hand, when receiver downloads or receives the data, can generate the MAC of received data and compare it with the MAC received along with the received data which was generated by the owner and if both the MAC codes are same then user is assured of the integrity of data, i.e., data has not been tampered.

4.5. Loss of user identity and password

For unauthorized access, authentication is required to be there in the cloud computing security structure. Thus, in case any user looses or by mistake reveals his user identity and password to any unauthorized person, the data can be in danger. To protect the data, we have added another parameter, which is a must to clear in order to access the data in cloud. Here the user will be asked a security question whose answer is known to the authorized user only, so the unauthorized user will face disappointment only even after having the correct user identity and password. Moreover, attacker must know the master key to decrypt the encrypted data received from the cloud.

5. Functionality analysis

An efficient cloud data security model should be able to overcome all the possible issues of cloud computing, so as to provide the benefits of cloud computing to reach its maximum heights and propel in the direction it is designed for, by preventing the owner's data from all the risks associated. Table 1 shows the comparison of the proposed model with other data security models.

Table 1
Functionality comparison.

	Juels et al. (2007)	Chor et al. (1995)	Wang et al. (2009)	Prasad et al. (2011)	Proposed model
Identification and authentication	Yes	Yes	Yes	Yes	Yes
Authorization	Yes	Yes	Yes	Yes	Yes
Confidentiality	No	No	Yes	Yes	Yes
Non-repudiation	No	No	Yes	Yes	Yes
Integrity	Yes	No	Yes	Yes	Yes
Encryption	Yes	No	No	No	Yes
Storage provider verification	No	No	No	No	Yes
Secure even after loss of user identity and password	No	No	No	No	Yes
Indexing of data	No	No	No	No	Yes
Keyword search	No	No	No	No	Yes

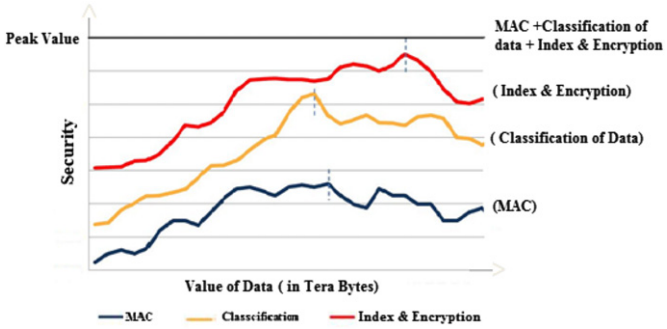

Fig. 10. Security evaluation.

The comparison of proposed model shows that it covers most of the possible security concerns by providing functions and techniques that are good enough to deal with the security issues of cloud computing. The following points are also used to keep the security of data.

1. Background screening of any employee or contractor who has responsibility for back-up or restore of data, if something is found, he must be prosecuted.
2. Right to pursue permissible legal action against any supplier employee or contractor's wrong doing if found by the cloud supplier or audit team.
3. Notification, within the context of applicable law (state/federal), of legal department for any confirmed breach into owner's data (query or removal).
4. Cloud computing supplier maintains security monitoring logs of all access to owner's data and documents access as routine, random audit, or suspicious leveraging their prescribed scripts and operational procedures as the basis for all audit, for not less than seven years.
5. Off-site back-up for disaster recovery and business continuity must be encrypted and all vendors must subscribe to all security measures above, without exception, including the audit.

6. Experimental results

The proposed technique is analyzed with respect to implementation. This model is tested on cloud computing simulator named Hadoop. Fig. 10 shows that the status of security after implementation of security parameters namely MAC, Classification of data and Index and Encryption technique. MAC provides less security than classification of data and classification of data provide lesser security than the implementation of index and encryption technique. Overall, the security of data related to owner is very good if combination of all three security parameters viz: MAC, Classification of data and Index and Encryption in taken into account. It results in very good security of the proposed model, which is represented as a peak value shown in Fig. 10.

7. Conclusion

The proposed technique provides a way to protect the data, check the integrity and authentication by following the best possible industry mechanisms. It introduces the division of data into different sections, Index builder,128-bit SSL encryption, Message authenticate code and a double authentication of user one by owner and other by cloud and verification of digital signature of the owner. It provides availability of data by surpassing many issues like data leakage, tampering of data and unauthorized access even from the cloud service provider. Proposed method achieves the availability, reliability and integrity of data traversing through owner to cloud and cloud to user. In addition to this, it also provides more flexibility and capability to meet the new demand of today's complex and diverse network and also enable the user to retrieve files from cloud by searching over an encrypted data.

Acknowledgment

This work is supported by University Grant Commission, India under major research project entitled "Trust based proactive resource provisioning in cloud" wide letter no. 40-255/2011 dated 29 June, 2011.

References

Bowers KD, Juels A, Oprea A. Proofs of retrievability: theory and implementation, Cryptology e-Print Archive. Report 2008/175; 2008a.
Bowers KD, Juels A, Oprea A. HAIL: a high-availability and integrity layer for cloud storage, Cryptology e-Print Archive. Report 2008/489, 2008b.
Cachin C, Micali S, Stadler M. Computationally private information retrieval with polylogarithmic communication, LNCS Springer Verlag, Advances in Cryptology- EUROCRYPT'99, 1592, p. 402–414, 1999.
Chor B, Gilboa N, Naor M. Private information retrieval by keywords. Report 98-03. Theory of Cryptography Library, 1998.
Chor B, Goldreich O, Kushilevitz E, Sudan M. Private information retrieval, In Proceedings of the 36th annual symposium on foundations of computer science, IEEE, p. 41–51, 1995.
Daniel EM, Wilson HN. The role of dynamic capabilities in e-business transformation. European Journal of Information Systems 2003;4(12):282–96.
Dikaiakos MD, Katsaros D, Pallis G, Vakali A, Mehra P. Cloud computing. IEEE Internet Computing 2009;12(5):10–3.
Gertner Y, Ishai Y, Kushilevitz E. Protecting data privacy in private information retrieval schemes, In Proceedings of the 30th annual ACM symposium on theory of computing, ACM, p. 151–160, 1998.
Juels A, Burton J, Kaliski S. PORs: proofs of retrievability for large files. Proceedings of CCS '07, p. 584–597, 2007.

Julisch K, Hall M. Security and control in the cloud. Information Security Journal: A Global Perspective 2010;19(6):299–309.
Kamara S, Lauter K. Cryptographic cloud storage. Lecture Notes in Computer Science 2010;6054:136–49.
Kusilevitz E, Ostrovsky R. Replication is not needed: single database, computationally-private information retrieval, In Proceedings of the 38th annual symposium on foundations of computer science, IEEE, p. 364–373, 1997.
Overby E, Bharadwaj A, Sambamurthy V. Enterprise agility and the enabling role of information technology. European Journal of Information Systems 2006;15(3):120–31.
Prasad P, Ojha B, Shahi RR, Lal R. 3-dimensional security in cloud computing. Computer Research and Development (ICCRD) 2011;3:198–208.
Popa RA, lorch JR, Molnar D, Wang HJ, Zhuang L , Enabling security in cloud storage SLAs with cloudproof. Technical report. Microsoft Research May 2010.
Shacham H, Waters B. Compact Proofs of Retrievability, Proceedings of Asiacrypt '08, 5350, p. 90–107, 2008.
Sood SK, Sarje AK, Singh K. A secure dynamic identity based authentication protocol for multi-server architecture. Journal of Network and Computer Applications 2011;34(2):609–18.
Wang C, Cao N, Li J, Ren K, Lou W. Secure ranked keyword search over encrypted cloud data. Journal of the ACM 2010;43(3):431–73.
Wang C, Wang Q, Ren K, Lou W. Ensuring data storage security in cloud computing, quality of service, 2009, IWQoS IEEE 17th international workshop, p. 1–9, 2009.