Efficient integrity checking technique for securing client data in cloud computing

Dalia Attas and Omar Batrafi

Computer Science Department

College of Computers and Information Technology

King Abdul Aziz University

Jeddah, Saudi Arabia

2011

Abstract: *It has been widely observed that the concept of cloud computing is become one of major theory in the world of IT industry. It involves storing the user's data to be able to use the applications and services that the clouds introduce. There is a significant numbers of risks can be occurred. One of these risks that can attack the cloud computing is the integrity of the data stored in the cloud [2].*

In order to overcome the threat of integrity of the data, the user must be able to use the assistance of a Third Party Auditor (TPA). The TPA has an experience that clouds users does not have, and checks over the integrity that is difficult for the users to check. The user can handout the integrity checking mission to the TPA, in such a way that the TPA will not be able to manipulate with the client data with one way or another [5].

In this paper, we will introduce a model for the integrity checking over the cloud computing with the support of the TPA using digital signature technique. The proposed model result was shown efficiently with a number of situations that performed by unauthorized attackers. The checking done over two parts the CSP and TPA, without giving any secure data that void the integrity definition and without uploading any secure data to the cloud.

Keywords: *data centers, Third party editors, cloud server, digital signature and encryption.*

1. Introduction

According to [1] "*Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction*".

In each environment there are some concerns that affect the usage and the behavior [2], below some of these concerns in the cloud:

- *Access:* When there is an unauthorized access to the data, the ability of altering on the client data arise.
- *Availability:* The data must be available all the time for the clients without having problems that affect the storage and lead to the client data lose.
- *Network Load:* The over load capacity on the cloud may drop the system out according to the high amount of data between the computers and the servers.
- *Integrity:* The data correctness, legality and security is the most fields that influence on the cloud and have major lay on the service provider.
- *Data Location:* The client does not know the actual place that the data saved or centered in because it distributed over many places that led to confusion.

One of the major concerns in the cloud computing is the data integrity, according to [3] to conserve the integrity of a specific point it must be defined, exact, changed suitably, changed by allowable peoples and process, dependable, and having an important effect. From above reference, in the cloud computing we need to ensure the integrity by making the user

capable to check over the cloud data from any unauthorized modification.

There are a lot of attackers that violate data integrity concept in the cloud [4], they can be attackers from inside the cloud environment for instance suspicious employee at CSP. The CSP is responsible for storage infrastructure and web services interface that can be used for storing and checking the user data. In addition, attackers can be from outside the cloud environment like the intruders or network attackers.

According to [5], we can check over the data integrity by enabling a new role which is TPA because it possesses experience capabilities that the customer does not. Auditors can understand the threats and they know best practices. Also they have the resources to check for process adherence and service quality. The TPA will be able to verify over any threats in online storage services that are represented in the cloud server. Thus, the user who owns the data can rely on the TPA to verify his\her data in the cloud without involving with the procedure.

The encryption idea is based on scrambling the information that only the one who have the secret key can expose it by decryption [6]. The encryption concept will not be enough to ensure the data integrity over the cloud. So in order to do the verifying over the CSP and the TPA, it would be preferable to use the digital signature technique because it does not revile any of the security parameters that violate the integrity concept [4].

In this paper we assume that the data integrity in the cloud is well verified using the TPA and digital signature technique and we expect the system to reach a fine grade of data validity and quality.

The rest of the paper organized as follows. Section 2 introduces the literature review. Then section 3 includes the proposed model which contains system model and implementation. Section 4 includes the results and evaluation, followed by section 5 which gives the conclusions; finally section 6 talks about the scope and limitation.

2. Literature Review:

Frequently, when we mention the cloud computing issue, enormous threats are raised .One of the major threats are data privacy and integrity. A lot of research discuss this problem and introduce many solutions to decrease the threat of the data privacy and integrity. Priya Metri and Geeta Sarote [4] introduce threat model to treat the privacy problem in the clouds. One of the threats in cloud computing is tampering with data in the cloud that interfere with the unauthorized modifications for the data, which lead to an effectiveness on processors, data storage and data flow. Then, they suggested different solutions technique for this threat. One of the solutions is using digital signature which will be used in our model.

Lately, a lot of researches moves into the concept of remotely stored data verification Ateniese et al. [7] have proposed a PDP (Provable Data Possession) model for verifying if an untrusted server stores a client's data. As shown in figure 1 , the data owner processes the data file to generate some metadata to store it locally. The file is then sent to be store in the server, and the owner may delete the local copy of the file. Clients may encrypt the file earlier to upload it to the storage. After that, the client asks the server to reduce a proof of the stored file, which it returned to the client. For security issues, the client verifies the "yes" or "no" answer from the server to make sure from his behavior.

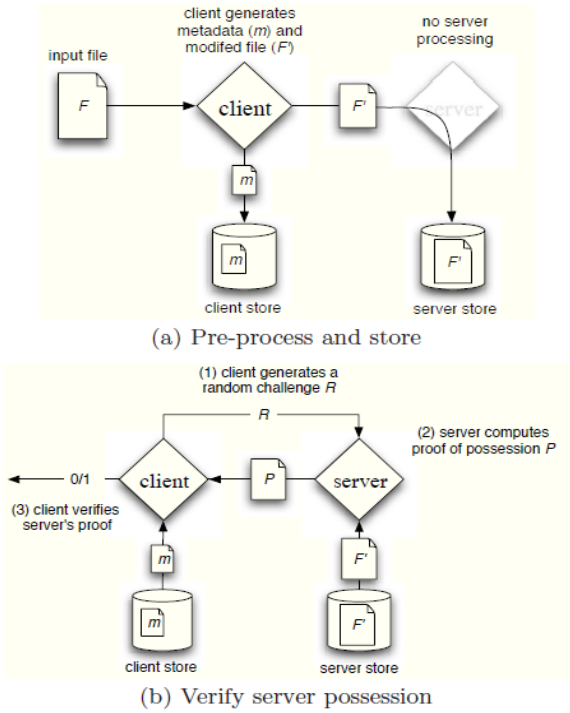

Figure 1. Protocol for Provable Data Possession [7]

The PDP scheme works in two stages, setup stage and challenge stage. In the setup stage, the client starts by generating the public key and the secret key, and then the client computes the tags for each file block and stores it at the server. The client then sends the public key and the file to the server for storage and deletes the file from its local storage. In challenge stage, the client requests from the server a proof of possession for a subset of the blocks in the file. Then the client checks the validity of the proof.

One of the adaptation tools for creating an online service is third party auditing because it allows customers to evaluate risks, and it increases the efficiency of risk reduction insurance. Mehul A. Shah et al. [5] proposed a solution to allow the clients to become aware of any changes to the stored data in the

server. By their solution, TPA can verify the integrity of files stored by a remote server without knowing any of the file contents. Their solution has three phases: initialization, verification, and extraction. In initialization phase, the storage service stores the document on behalf of the client, and the auditor initializes long-term state. During verification phase, the auditor frequently checks the stored data. At the end, in extraction phase, the auditor will return the data to the customer in case of distrust.

The main purpose of the solution planned is to make sure that the remote server properly possesses the client's file, and to prevent any information outflow to the TPA which is responsible for the verifying mission. Hence clients can way out to the third party editor to check the integrity of outsourced data, and this third party auditing process should not bring on any manipulation on the client's data. Besides of the auditing process, TPA involves with the server to check the stored data is undamaged and to extract of digital contents.

3. Proposed Model

- **System Model**

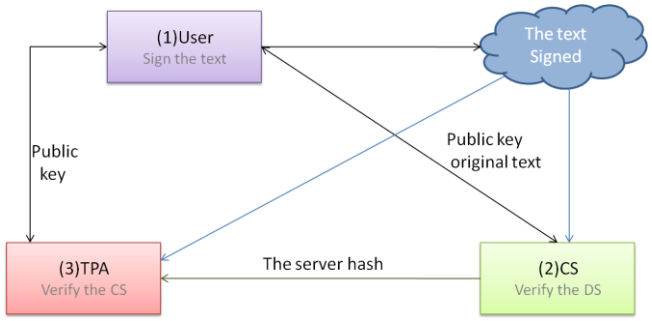

Figure 2. The Proposed Architecture

In Figure 2, the proposed architecture is introduced to provide data integrity in could system which has three main roles as below:

1- User: who uses and relay on the verification of the cloud to store a large data files with a secure way.

2- Cloud Server (CS): who is managed by CSP, which has a huge storing space and flexible recourses to keep up the client data.

3- TPA: who has the skills that clouds users does not have, and checks over the integrity that is difficult for the user.

When we upload a large data file in a remote cloud server, the users will not be able to process their data in local; moreover, they cannot verify their data in the remote servers. In that case they need to assign the TPA to do the verifying task in order to help the user in the verification that is difficult procedure and no any regular client can perform it without having a background in the verification process [5]. The TPA has an access to the cloud provider environment and understands the service level agreements (SLA) that is between the customer and the provider. By this way the TPA is reliable, and independent [5].

3.2. Implementation

To implement our design, we need to achieve some goals in our model by allowing the TPA to verify the correctness over the cloud data. Additionally, we need to ensure that the cloud server does not manipulate or alter the user data in the cloud.

In our construction, we consider on going the PDP model [7] that is suitable for verifying over the untrusted servers who stores a client's data. Furthermore, the model is achieved using the digital signature technique.

The digital signature works by taking the user data first, then perform a hash function over it using Message-Digest Algorithm (MD5). After that, computes the signature for the generated hash value by encrypting it with the private key. In the other side, the decryption is done by the public key but the result will be a hash value, and the hash value is not reversible to its original data.

There are three procedures in our model to satisfy the integrity concept:

1- Digital signature part will be done by the user.

2- The CS verifies over the user data in the cloud to check over the manipulation or intrusions in the cloud data.

3- The TPA verifies over the cloud server part to check if the cloud server was manipulating in the user data or not.

In next the paragraphs explanation for each entity function in the proposed model:

1- User: User first chose a random parameter to construct the public and the private keys then he\she will sign the data using the private key to be uploaded to the cloud, then he\she send the signed data to the cloud server and deletes its local copy.

2- CS: CS will compute a hash value from the original data to send it to the TPA, and then takes this hash value along with the data signed in the cloud for verification using the public key. At the end, the CS will inform the user if the data in the cloud intruded or not.

3- TPA: After the cloud server finishes its role, the TPA will be initiated to verify over the cloud server work by taking the hash value from the cloud server. TPA will take the data signed form the cloud and decrypt it with the public key. The

decryption will result a hash value that will be compared along with the hash value that the cloud server compute it in his part. After finishing the verification, the TPA will inform the user if the CS was trusted or not.

4. Results and Evaluation:

The objectives of achieving the integrity in the cloud was concurring in first place by checking over the client data from any attacking occur either was from outside the cloud like intruders or inside like the cloud server. In our paper we check over two attackers: inside and outside attacker. In case of the outside attacker the cloud server will checks over the client data integrity, and in case of the inside attacker the TPA will checks the cloud server integrity.

In order to evaluate our work, we need to experiment our model over all the situations that will happen if an attacker harm the system as shown below:

- *Situation 1:* The cloud data was not corrupted by either outside attacker or inside attacker (Figure 3). In this situation the user will get the success message from both the CS and TPA.

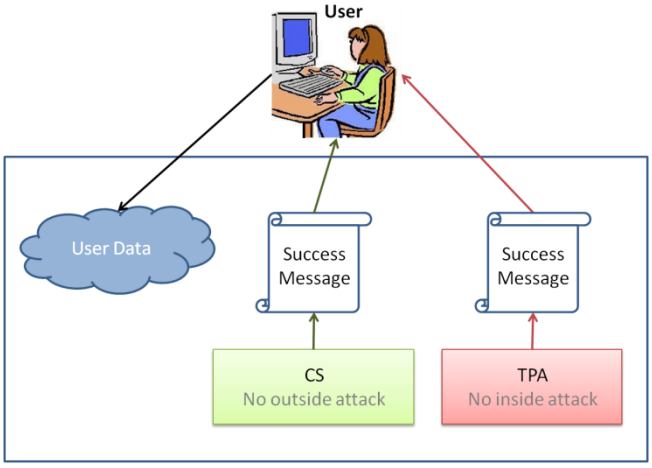

Figure 3. No Inside or outside attack

- *Situation 2:* The cloud data was corrupted by outside attacker (Figure 4)

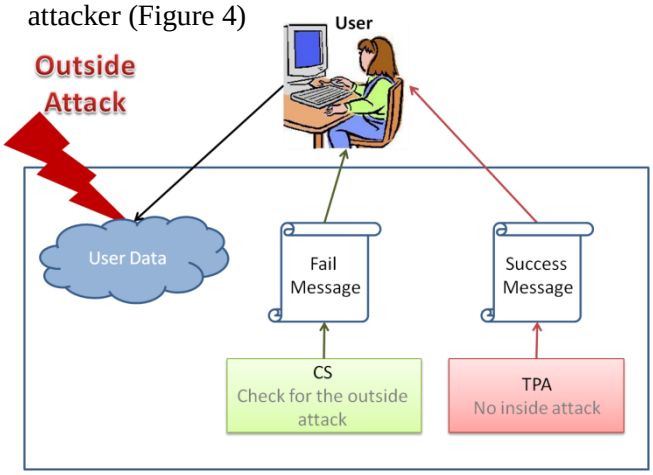

Figure 4. The Outside attack

If an attack occur from outside the cloud, the CS is responsible to check over this attack. The CS first check over the cloud by taking the user data and sign it to compare the signed value with the cloud data. The CS will find the compared value is not matched which means the data in the cloud corrupted. So CS will announce to the user the fails massage.

- *Situation 3:* The cloud data was corrupted by inside attacker (Figure 5)

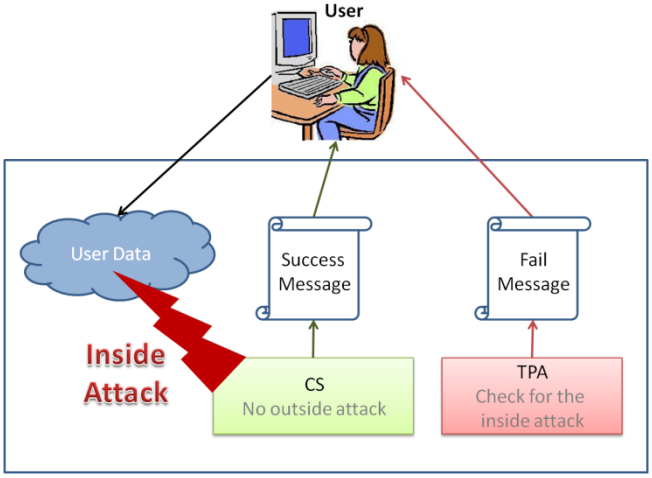

Figure 5. The Inside attack

If an attack occurs inside the cloud environment, the TPA will check over this attack by decrypt the cloud data and match it with the hash value from the CS. So the TPA will find that the match process fails and announce to the user the fails massage.

- *Situation 4:* The cloud data was corrupted by inside attacker and outside attacker (Figure 6)

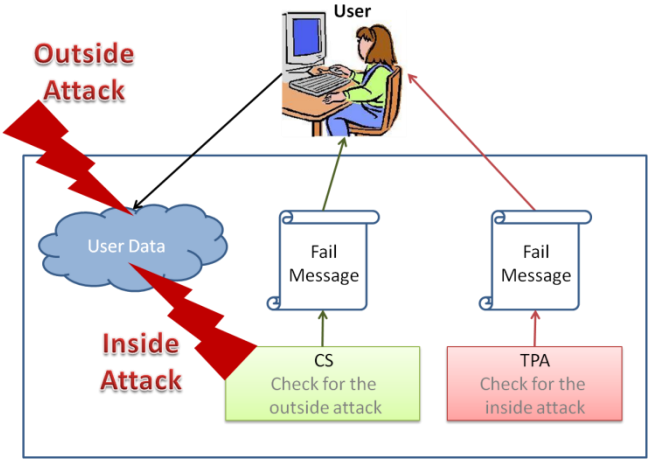

Figure 6. The Inside and Outside attack

If an attack occurs from inside and outside, the CS will check over the outside attack by checking the user data after signing with the signature of the cloud data, and the TPA will check over the inside attack by matching the decrypt value from the cloud data with the hash value from the CS. The CS and the TPA will find that the matching process fails and then inform the user by the fail massage.

The model was implemented by building the project using Visual Studio 2010 and Windows Azure tool to run the cloud environment that support creating web application in local machine. The languages used in programming the system are C Sharp with ASP.NET in order to create three web pages each concern with a specific job in the model: the user, CS and TPA.

In Figure 7 the user page have two input boxes, one input for the random parameter and the other for the text to be signed. Additionally, one outbox add for the text after digitally signed. The user has the sign button to sign the text and upload button to upload the signed text to cloud.

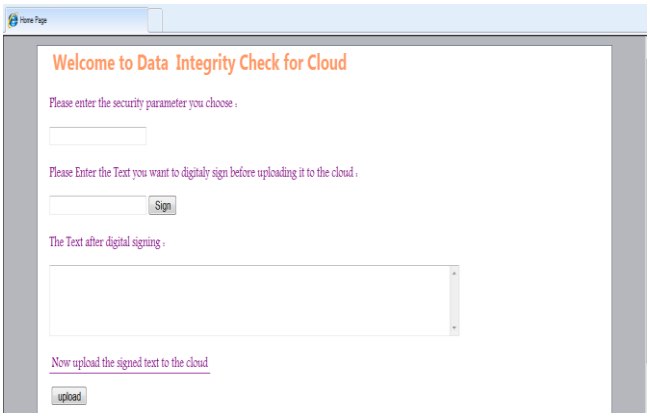

Figure 7. User Page

In Figure 8 the CS page have two buttons: one to check by the CS to verify over the user data in the cloud and display "yes" or "no" depending on the answer from the CS. The other button is to check by the TPA that will transfer the user to the TPA page that is demonstrated in (Figure 9). In order to go to the TPA page the user must first check by the CS button to make the TPA able to verify over the CS.

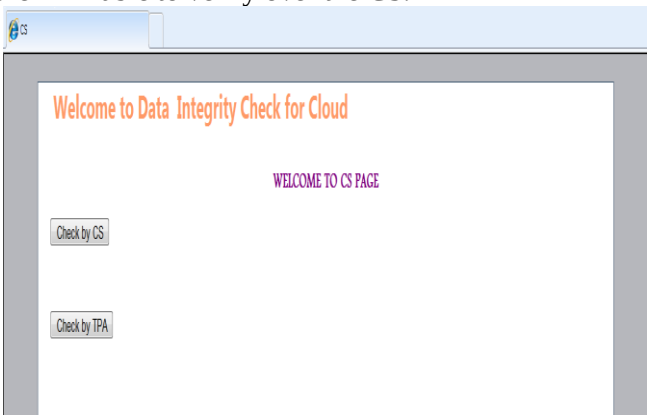

Figure 8. CS Page

In Figure 9 TPA page have a button to check by the TPA over the CS which will return the choice "yes" or "no" depending on the checking result over the CS integrity.

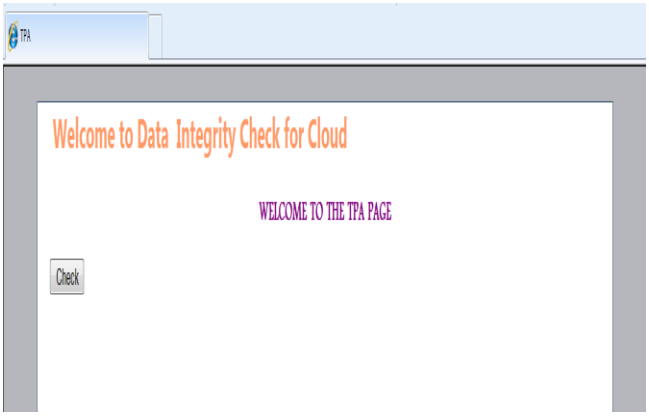

Figure 9. TPA Page

Consequently, we found that our model provides us a well-defined and efficient result depends on each situation we assumed. With the pages which have been created, we make the system easy to use by the user that does not have background in the digital sign technique.

After discussing the situations for the attacking as shown earlier and the forms that were built using windows Azure application make the cloud able to reach the client data integrity and availability.

5. Conclusion:

In this paper, we proposed a model for the integrity check over the cloud computing and we utilize the TPA and digital signature to achieve the integrity concept, in such a way to help the user to verify and examine the data from unauthorized people that manipulate with the cloud or extract from the data. Moreover, we are able to evaluate our work using a windows azure project that involves digital signature coding. As results, we found that our model worked well according to our claims.

6. Scope and limitation :

In our paper we decided to concern about the client data storing service in the cloud. The main objective was to study the ability to verify the client data with the absent of the editing and the deleting. As part of the verification process we assume that the TPA is reliable and independent according to the service level agreements (SLA), which does not mean that there is no space for the TPA to cheat. The approach used for the encryption in the verification process was the digital signature. In the implementation we used as an example of the client data in the cloud a text entered by the client, this research is not covering other various kinds of client data.

Future work

In future work there is a capacity for verifying over the other services that the cloud performs. It would be preferable to check over the ability to edit or delete the data in the cloud. In addition, TPA would be enhanced if there is a serious proof for its credibility. The verification process could be done by further techniques rather than the digital signature to improve the efficiency and security. We can improve the implementation by enabling the user to enter different kinds of data to be verified.

Acknowledgements

I would like to show my gratitude to Dr. Omar Batrafi whose encourage me from the start to the final level to build this fine piece of paper.

Lastly, it is my pleasure to thank my parents, my husband, and my child whom made this paper possible.

References

[1] Mell P. and Grance G., "The NIST Definition of Cloud Computing (Draft)," *in Proceedings of the National Institute of Standards and Technology*, Gaithersburg, pp. 6, 2011.

[2] Chandran S. and Angepat M., "Cloud Computing: Analyzing the risks involved in cloud computing environments," *in Proceedings of Natural Sciences and Engineering*, Sweden, pp. 2-4, 2010.

[3] Pfleeger C. and Pfleeger S., *Security in Computing*, 2nd ed, Prentice Hall, New Jersey, 1997.

[4] Metri P. and Sarote G., "Privacy Issues and Challenges in Cloud computing," *International Journal of Advanced Engineering Sciences and Technologies*, vol. 5, no. 1, pp. 5-6, 2011.

[5] Shah M., et al., "Auditing to keep online storage services honest" *in Proceedings of HotOS'07*, Berkeley, CA, USA, pp. 1-5, 2007.

[6] Kaufman C., Perlman R., and Speciner M., *Network Security: Private Communication in a Public World*, 2nd ed, Prentice Hall PTR, New Jersey, 2002.

[7] Ateniese G., et al., "Provable data possession at untrusted stores," *in Proceedings of CCS'07*, New York, USA, pp. 598-603, 2007.