Framework For Secure Cloud Data Communication

Nitika Aggarwal, Abhishek Choudhary, Maalvika Bachani, Mrs. Rachna Jain

Abstract: Cloud computing is an emerging technology where all the computing resources are shared on the cloud rather than having local servers or personal devices to handle applications. Encrypting data residing on the cloud database is required to prevent unauthorized access of confidential and critical information and the subsequent modification of the information for personal benefit. As all the organizational information resides on the computers, security of this data is of utmost importance. Shamir's secret sharing algorithm is one successful way of encrypting the data. In this research paper we have developed a new encrypting algorithm based on symmetric key cryptography. We have used logical operations like XOR and zero padding. The Shamir's secret sharing algorithm would act on the key generated by our proposed encryption algorithm. This algorithm is an efficient and a simple strategy for secure communication in cloud computing.

Index Terms: Algorithms, Cloud Computing, Ciphertext, Decryption, Encryption, Lagrange's basis polynomial equation, Security, Shamir's secret sharing algorithm, Threshold value, Zero Padding

1 INTRODUCTION

Cloud Computing refers to manipulating, configuring and accessing the applications online. It offers online data storage, infrastructure and application. Cloud computing is an internet-based computing in which large group of remote servers are networked to allow sharing of data-processing tasks, centralized data storage, online access to computer services or resources. Cloud computing relies on sharing of resources to achieve coherence and of scale similar to a unity over a network. Clouds can be classified according to the deployment models as public, private or hybrid. Public cloud: The public cloud allows systems and services to be easily accessible to the general public. Public cloud may be less secure because of its openness e.g e-mail. Private Cloud:The private cloud allows systems and services to be accessible within an organization. It offers increased security because of its private nature. Hybrid Cloud:The hybrid cloud is a mixture of public and private cloud. However, the critical activities are performed using a private cloud while the non-critical activities are performed using public cloud. Further the cloud can be classified on the basis of the services it provides. IaaS (Infrastructure as a service), PaaS(Platform as a service), SaaS(Software as a service).

- *NitikaAggarwal is currently pursuing Bachelors of technology(B.Tech) in BharatiVidyapeeth's College Of Engineering, IPU, India, PH-9958027773.*
  E-mail: nitika.aggarwal93@gmail.com
- *Abhishek Choudharyis currently pursuing Bachelors of technology(B.Tech) in BharatiVidyapeeth's College Of Engineering, IPU, India, PH-9711740899.*
  E-mail: abhishek.0909ch@gmail.com
- *MaalvikaBachaniis currently pursuing Bachelors of technology(B.Tech) in BharatiVidyapeeth's College Of Engineering, IPU, India, PH-9717134438*
  E-mail: maalvika.bachani@gmail.com
- *Mrs. Rachna Jain is the assistant professor at BharatiVidyapeeth's College Of Engineering, IPU, India.*
  E-mail: rachna.jain@bharatividyapeeth.edu

Infrastructure as a service (IaaS): This is the most basic level of service. Each of the service models make use of the underlying service model i.e each inherits thesecurity and management mechanism from the underlying model.IaaS provides access to fundamental resources such as physical machines, virtual machines, virtual storage, etc. Platform as a service(PaaS):PaaS provides the runtime environment for applications, development and deployment tools, etc. Software as a service(SaaS):SaaS model allows to use software applications as a service to end users. Microsoft Azure is a cloud computing platform and infrastructure for building, deploying and managing applications and services through a global network of Microsoft-managed data centers. It provides both PaaS (Platform as a service) and IaaS (Infrastructure as a service) services and supports many different programming languages, tools and frameworks. We have used Microsoft Azure to deploy cloud and have used its services. Next the encryption algorithm designed is applied to the data that is sent to the cloud and then it is decrypted using the inverse algorithm during the retrieval of the information from the cloud.

2 NEED FOR SECURITY IN CLOUD

With Cloud Computing rapidly gaining popularity, it is important to highlight the resulting risks. Security and Privacy is the biggest concern about cloud computing. Since data management and infrastructure management in cloud is provided by third-party, it is always a risk to handover the sensitive information to such providers. Although cloud computing vendors ensure more secure password protected accounts, any sign of security breach would result in loss of clients and business. On the cloud as all the data resides at the same place if ever there is a breach all the critical and sensitive information will leak.

3ENCRYPTION ALGORITHM DESIGN

3.1 Terms Used:

1. Plaintext: The original message produced by the sender i.e the data before encryption.
2. Ciphertext: The plaintext is transformed into ciphertext. The encryption algorithm converts the plaintext into ciphertext.
3. Decryption: Decryption is a process of transforming back the ciphertext beck to plain text.

4. Ciphers: The encryption and decryption algorithm together are referred to as ciphers.
5. Key: A key is a value or a number. The keys are secret.
6. Symmetric Key Cryptography:Symmetric key cryptography is also known as the secret key cryptography. In the symmetric key cryptography, the same key (shared secret key) is used by the sender and receiver. The sender uses this key alongwith the encryption algorithm to encrypt data, and the receiver uses the same key alongwith the decryption algorithm to decrypt the data. The encryption algorithm makes use of a combination of addition and multiplication whereas the decryption algorithm uses a combination of subtraction and division.

3.4 Functions Used

1. XOR:Exclusive OR or exclusive disjunction is a logical operation that outputs true whenever both the inputs differ (one is true, the other is false).
2. Padding:The primary use of padding is to dampen the predictability of cribs in any message. This way breaking of the encryption becomes difficult. The random length of the padding also prevents an attacker from knowing the exact length of the plaintext.

Example: if 12 45 98 10 36 44 57 41 are the tokens produced then the file will contain the following text:“120450980100360440570410”

3.5 Proposed Work

This algorithm is a two pass algorithm. The data to be encrypted is sent through two encryption passes and then the key that is generated is used to implement the Shamir’s secret sharing algorithm. This algorithm designed uses 32 tokens for encryption. These tokens can be anything ranging from numbers to characters. The tokens are separated by a #.

for e.g.—

Tokens used:
12019142g171148177ta19659991713077n4811959199U7969
1251941994519367p138146

Actual tokens for the key separated by a #.

120#19#142#g#171#148#177#t#a#196#59#99#171#30#77#n#48#119#59#199#U#79#69#125#194#199#45#193#67#p#138#146#

1) In the first pass the tokens in the key is XORedwith the characters in the plaintext. The first token is XORed with the first character, the second with the second character, the third with the third and so on. When the content of the file or plaintext is more than the length of the key then it starts from the first token of the key to the last token and then repeats the XORing the next character in the plaintext with the first token.
2) In the second pass all the tokens of the key are again XORed with the tokens of the cipher text of the first pass but in an opposite manner. The last token of the key XORed with the first token character of the ciphertext of the first pass, then the second last token of the key with the second token character of the cihertextf the first pass, and so on. This also repeats in a cyclic manner. The output obtained is now padded with a zero and then stored in a temporary file.

i.e. if 12 45 98 10 36 44 57 41 are the tokens produced then the file will contain the following text:“120450980100360440570410” After padding it with 0 the cipher text is stored in an opposite fashion i.e.“014075044063001089054021” and then it is stored on the local system/cloud.

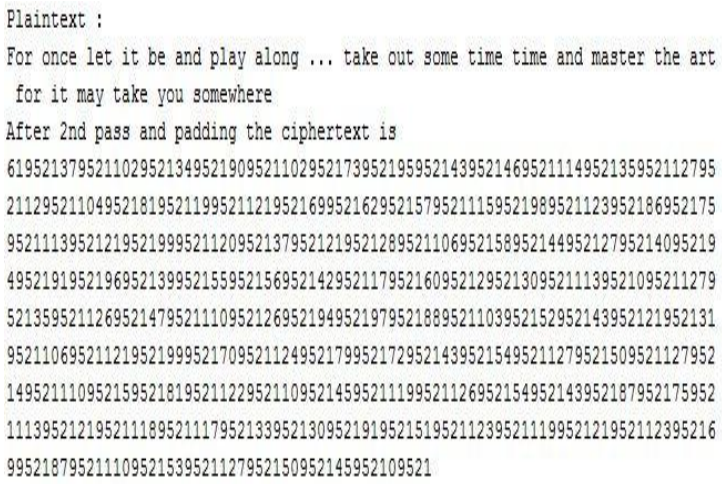

Figure 1: *The plaintext and the encrypted and decrypted outputs obtained after the second pass of the proposed algorithm.*

3) Shamir’s secret sharing algorithm: This secret sharing algorithm has been created by Adi Shamir. The secret is divided into small parts and each and each of the participant is given a unique part and during reconstruction all or some of these parts need to be retrieved and combined. Relying that all of the participants will be able to contribute during the re-construction of the secret is impractical therefore a threshold value (k) is selected and if the threshold number of participants contribute then the secret have be generated back.

The definition of Shamir’s algorithm states that Let S be the secret. We divide this secret into n smaller parts S1, S2,S3,…. Sn. This is done in such a way that when we have knowledge of any k or more Si pieces, the secret can be easily re-generated. When the knowledge is of k-1 or lesser Si then the secretremains undetermined. This s called a (k,n) threshold scheme. When k=n the all the participants are required to reconstruct the secret.The central idea is based to the concept that two points are sufficient to define a line, three points are sufficient to define a parabola, four points to define a cubic curveand so forth. Thus, it takes k points to define a polynomial of degree k-1. We use (k,n) threshold scheme to share the secret S. Choose at random k-1 positive integers a1,a2,…a(k-1) with a(i)<P(prme number), and let a0=S. The polynomial is then build f(x)=a0 + a1x +a2x^2 + a3x^3 + …… + a(k-1)x^(k-1). Constructed any n points out of it, for instance set i= 1,…n to retrieve (i,f(i)). Every

participant is given a point (an integer input to the polynomial, and the corresponding integer output). Given any subset of k of these pairs, we can find the coefficients of the polynomial using interpolation. The secret is the constant term a0.

Example:
. Let S=1234
We wish to divide the secret into 5 parts (n=5), where any subset of 3 parts (k=3) is sufficient to reconstruct the secret. At random we select (k-1) numbers: 144 and 94.

(a1=144; a2=94)

Our polynomial to produce secret shares (points) is therefore:

$$f(x) = 1234 + 144x + 94x^2 \quad (1)$$

We construct 5 points $D_{(x-1)} = (x, f(x))$ from the polynomial:

$$D_0 = (1,1472)$$
$$D_1 = (2,1898)$$
$$D_2 = (3,2512)$$
$$D_3 = (4,3314)$$
$$D_4 = (5,4304)$$

Here the key that is generated after the two passes of the encryption is then passed on to the Samir's secret sharing algorithm. It is further encrypted and stored.

4 DECRYPTION TECHNIQUE

1. In order to reconstruct the secret any 3 points will be enough.We will compute *Lagrange basis polynomials*:

$$L_0 = ((x - x_1)/(x_0 - x_1)) * ((x - x_2)/(x_0 - x_2)) \quad (2)$$

$$L_1 = ((x - x_0)/(x_1 - x_0)) * ((x - x_2)/(x_1 - x_2)) \quad (3)$$

$$L_2 = ((x - x_0)/(x_2 - x_0)) * ((x - x_1)/(x_2 - x_1)) \quad (4)$$

Therefore, $f(x) = \sum [y(i) * lj(x)]$ where j=0,1,2

$f(x) = 1234 + 144x + 94x^2$ .The secret is the free coefficient, which means that S=1234, and we are done.

2. Then the content of the file is fetched and rotated to get it in the correct order. "01407504406300108905402 1" to "12045098010036044057041 0"

3. Now remove the padded zeroes and get the tokens. Now again we have two passes of the decryption as well.

4. In pass one the tokens are fetched and are then XORed with the tokens of the key in an opposite fashion i.e the first token of the ciphertext with the last token of the key.

5. In the second pass the generated tokens from the first passare again XORed with the corresponding tokens of the key in a cyclic manner. If the length of the ciphertext is more than 32 then the first token of the key is used again.

6. The final tokens are then calculated by converting them into characters using their ASCII codes to get the plaintext.

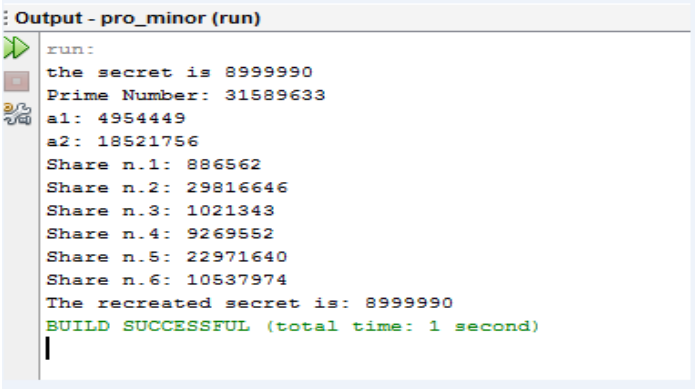

***Figure 2:** The output of the Shamir's secret sharing algorithm.*

5 CONCLUSION

Through this research we conclude that the risks on the security and privacy of data can be significantly improved by using the above proposed algorithm. The proposed algorithm is efficient and simple.

6 FUTURE SCOPE

This algorithm can be further improved upon by making it a four pass algorithm. This will increase the reliability of the encryption. In the third and fourth pass logical operations like XOR and shifting can be done in a cyclic manner and acyclic manner respectively.

7 REFERENCES

[1] Khalil Challita, HikmatFarhat. Combining Steganography and Cryptography: New Directions. In International Journal on New Computer Architectures and Their Applications (IJNCAA) 1(1): 199-208.The Society of Digital Information and Wireless Communications, 2011 (ISSN 2220-9085)

[2] http://www.tutorialspoint.com/cloud_computing/cloudcomputing_overview.htm

[3] Jianfeng Yang, Zhibin Chen. Cloud Computing Research and Security Issues.

[4] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, Andrew Konwinski, Gunho Lee, David A. Patterson, Ariel Rabkin, Ion Stoica, MateiZaharia. Above the Clouds: A Berkeley View of Cloud Computing. In Above the Clouds: A Berkeley View of Cloud Computing.

[5] G. Jai Arul Jose, C. Sajeev. Implementation of Data Security in Cloud Computing. In International Journal of P2P Network Trends and Technology- July to Aug Issue 2011.

[6] Lisa J. Sotto, Bridget C. Treacy, Melinda L. McLellan. Privacy and Data Security Risks in Cloud Computing.

[7] SeongHan Shin, KazukuniKobara. Towards Secure Cloud Storage. In demo for CloudCom2010

[8] Robert Denz, Stephen Taylor. A survey on securing the virtual cloud. In Journal Of Cloud Computing, a SpringerOpen Journal.