International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 02 Issue: 09 | Dec-2015 www.irjet.net p-ISSN: 2395-0072

Storage Space Minimization and Data Security using Multiple Key Cryptography

Swati Patil, Keyuri Pophale, Rashmita Salian, Madhavi Thakare

[1234] *Student , Computer Department, KKWIEER, Maharashtra,India*

Abstract - ***Data storage has become a major issue in this Internet driven world. Therefore it has become a necessity for today's rapid growing and changing data to optimize the available space. In this system a new methodology is being proposed for performing data to image conversion in order to reduce the space required to store the data. Another problem that we are facing today is data theft. It becomes very important to add security measures to the data that we are storing in different databases. Hence in this system, we are focusing on the principles of data security. Multiple key cryptography (MKC) technique that performs the process of cryptography by making use of multiple keys which provides security to the storage data is applied. The system will be able to show space minimization strategy in the storage using data-to-image-to- data conversion procedure. Owing to the insufficient space available with the system, most of the people opt for cloud storage services. But if every other organization will prefer to outsource their resources on cloud then this would again lead to storage problem of dynamically growing data. Thus by making use of this approach the size of the data can be reduced to certain extent so that the space required to store the data would be less than previously required. In addition to this the cost associated with the storage of data on cloud can also be minimized. This in turn helps reduce power consumption in cloud and thus serves the purpose of green computing.***

Key Words: *cost minimizing , cryptography , multiple key, storage , security*

1. INTRODUCTION

Next-generation technologies generate a large amount of data. Individual companies, of course, maybe able to deal with their own data-storage should be within budget, either by redistributing some of the storage resources to the cloud or simply endowing in more local infrastructure. But the worldwide rate of increase means that eventually HDDs will reach its practical limits as storage medium-even in the cloud environment. The two alternative approaches to the problem can be: phase out old or useless data to make up the difference or develop a new storage technology that provides a cost saving alternative approach to HDDs. Simply waiting for SSDs to fall in price may not cut it. Simply deleting unneeded data is a tempting option-until you give it some careful thought. Obviously, the significance of the data does not rely on its age; some older data may be far more valuable than the newer data. Some reasonable criteria can be determined; this can be done with the help of certain algorithms. Data storage minimization techniques are applied and using Multiple key Cryptography we can assure good security to the files.

2. LITERARY SURVEY

- The data size reduction using some compression techniques may be one of the solutions to reduce power cost. But decompression is required during accessing this compressed data. So it will increase computational time during decompression which will consume more power than uncompressed data access.

- Software based encryption policy may also be adopted to secure the data which are very easy to implement. On the contrary, it will also increase computational time during accessing that compressed data. So there must be a trade-off among energy, data space, computational time and data security for minimizing the cost.

- Three different procedures to improve data center efficiency so that its cost can be minimized were identified by Greenberg et al.[2]. Their procedures include increase data center network agility, pursue design algorithm and market mechanism for optimizing resource usage and finally geo-diversified data center for performance improvement and reliability. But

they did not mention any statistical data about the data center cost minimization i.e. how much cost can be reduced? Besides, their three methods are not easy to implement in practice by the cloud service providers.

- Nicolae[3] applied adaptive transparent data compression technique for reducing the storage space and bandwidth used with a slight computational overhead. However the compression technique failed to compress multimedia data such as audio, video and image. Therefore, in our research we show how cost minimization and data security can be ensured using encryption over cloud data.

3. COST MINIMIZATION PROCEDURES

This approach reduces the space required to save data imparting security measures to the data, so we can say that cost minimization is possible through less space used. The proposed method is overcoming the drawbacks of the earlier solutions and provides substantial security using multiple key cryptography (MKC). Finally we can conclude that we can conduct more researches in future to minimize the storage space required for big data.

3.1 Multiple Key Cryptography

Multiple encryption [1] is the process of converting an original message into an unreadable form by performing encryption number of times, either by implementing the same or different algorithm strategies. It can also be referred as cascade encryption, cascade ciphering, multiple encryption and super encipherment. We are aware that symmetric cryptography is done with the help of single key encryption and public key cryptography which is performed through two keys named private key and public key. Multiple key cryptography is such a process where encryption and decryption is done through multiple keys. So multiple key encryption is used to convert plaintext into cipher text. Here same multiple keys are used to convert cipher text in reverse order to find the main plaintext during decryption which is called multiple key decryption (MKD).

3.2 Random Key Generator

Key generation is the process of generating keys for cryptography. A key is used to encrypt or decrypt whatever data is being encrypted/ decrypted. Computer cryptography makes use of integers for generating the keys. In some cases keys are randomly generated using a Random Number Generator (RNG)[1]. A random number generator (RNG) is a computational or physical device or a piece of software code which is designed to generate a sequence of numbers or symbols that cannot be reasonably predicted better than by a random chance. Here the input range for generating every random number depends on bit length of the keys and nature of data and operation to be performed on data. For example- if n-bit additive key is needed to be generated then input range will be 20 to 2n-1. In this project we need five different keys.

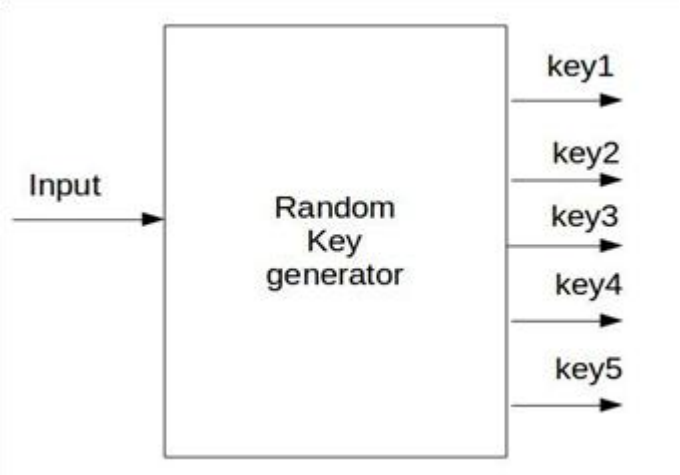

Fig -1: Random key generator

4. DATA SECURITY APPROACHES

Data security is a challenging issue. The goal of data security is to keep integrity, confidentiality, and availability of data. So it is the responsibility of both researchers and developers to ensure data security. A number of techniques are proposed by different researchers all over the world. Some of these techniques are shown in the following subsections

4.1 Data Hiding Based Security

Data hiding is the process of securing data from unauthorized access. There are several data hiding techniques such as cryptography, hashing, steganography, etc. It is very challenging to provide data security through steganography and cryptography because of increasing computation time. So there should be a standard trade-off among security, encryption, cryptography, and computational overhead i.e. how much computational overhead can be tolerated for security? In 2010, Kamara et al[4] demonstrated the complete architecture of cryptographic cloud storage which will be advantageous for both customers and service providers in terms of their data security. But they have implemented the process of encryption and decryption using unique keys which are less secured than MKC policy. In 2012, Gampala et al[5]

explored data security of cloud by implementing digital signature and encryption with elliptic curve cryptography. But they did not show the security measurement of their method when it will be applied to major varieties of data in the cloud.

4.2 Data Fragmentation

In computer storage, fragmentation can be defined as the phenomena in which the storage space of the system is not used efficiently, reducing the capacity, the performance or may be both in worst case scenario. Fragmentation in many cases can lead to wastage of space which can degrade the performance of the system.

4.3 Data Centric Security

The purpose of data centric security targets at directly providing protection to the data. This will increase effectiveness to the security measures currently in place. This solution provides security throughout its lifecycle.

`

5. OVERVIEW OF ALGORITHM

5.1 Encryption

Initially we will accept the file to upload from the user. Then we will read data byte by byte from the file. Then we will calculate the size of the data and divide the data into 3 equal vectors. Using Random Key Generator, generate 5 random keys. Then encrypt the 3 vectors using these keys. The resulting Encrypted RGB image is stored in the disk.

5.2 Decryption

The data which is encrypted by using MKC technique can be regenerated i.e.it can be decrypted by using the previously generated five different keys but in a reverse fashion. The encrypted RGB image is divided into 3 parts and then decrypted using the same keys which were used during encryption. The re-generated file is then given back to the user.

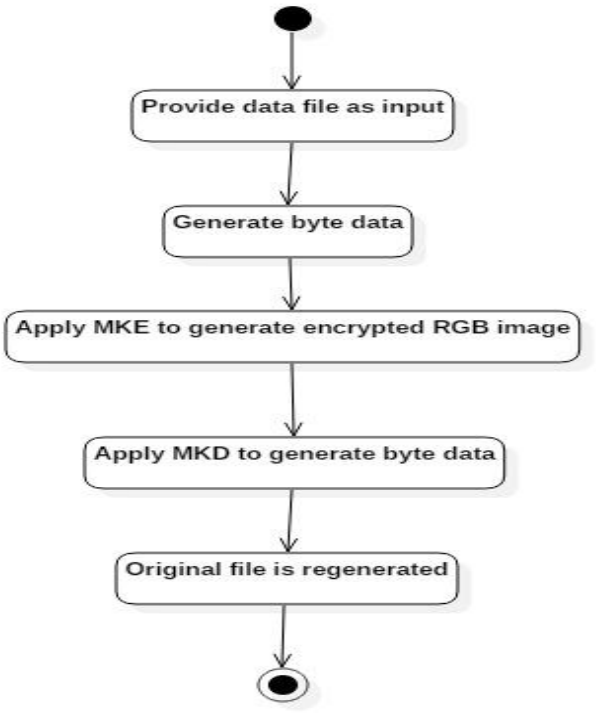

Fig -2: System flow diagram

6. FUTURE SCOPE

Processing big data is a concern for researchers. They can work on this algorithm to minimize the computational overhead using distributed processing on a machine with standard RAM and processor for processing big data in the cloud. This algorithm fails to compress image file with JPEG extension and zip file where more research is needed to reduce these kinds of files. A limited number of files are evaluated through this algorithm so we can evaluate its performance using hundred types of files in future.

REFERENCES

[1] Tushar Kanti Saha, A B M Shawkat Ali,(2014). Storage Cost Minimizing in Cloud – A Proposed Novel Approach Based on Multiple Key Cryptography.

[2] A. Greenberg, J. Hamilton, D. A. Maltz and P. Patel, "The cost of a cloud: research problems in data center networks." ACM SIGCOMM Computer Communication Review 39.1 pp. 68-73, 2008.

[3] B. Nicolae, "High throughput data-compression for cloud storage." Data Management in Grid and Peer-to-Peer Systems. Springer Berlin Heidelberg, pp. 1-12, 2010.

[4] S. Kamara and K. Lauter. "Cryptographic cloud storage." Financial Cryptography and Data Security. Springer Berlin Heidelberg, 2010. 136- 149.

[5] V. Gampala, S. Inuganti and S. Muppidi, "Data Security in Cloud Computing with Elliptic Curve Cryptography." International Journal of Soft Computing and Engineering (IJSCE) pp. 2231-2307, 2012.