Sudan University of Science & Technology
College of Graduate Studies
Computer science

Confidentiality of Data in Public Cloud Storage Using Hybrid Encryption Algorithms

سرية البيانات في المخزن السحابي العام بإستخدام خوارزميات التشفير الهجين

A thesis submitted for the partial fulfillment of the requirements of M.Sc. degree in computer sciences

Submitted By
Khabab Mustafa Al-Khalifa Mohammed

**Supervisor **

D Faisal Mohammed

January- 2019

الآيـــــــــــــــة

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

﴿ تَبَارَكَ الَّذِي نَزَّلَ الْفُرْقَانَ عَلَى عَبْدِهِ لِيَكُونَ لِلْعَالَمِينَ نَذِيرًا (1) الَّذِي لَهُ مُلْكُ السَّمَاوَاتِ وَالْأَرْضِ وَلَمْ يَتَّخِذْ وَلَدًا وَلَمْ يَكُنْ لَهُ شَرِيكٌ فِي الْمُلْكِ وَخَلَقَ كُلَّ شَيْءٍ فَقَدَّرَهُ تَقْدِيرًا (2)﴾ **صدق الله العظيم**

[سورة الفرقان الاية 1-2]

Dedication

This thesis is dedicated to my parents, Basamat and Mustafa, who have always loved me unconditionally and whose good examples have taught me to work hard for the things are inspiring to achieve. This thesis is also dedicated to my best friends, Abdelaziz, Ahmed, Mojahed, Mohammed, Shash, Khalid, Omer, Sadam, Lotfi who have always been a constant source ofsupport and encouragement during the challenges of my whole collegelife. Also,to my Brothers, Sisters and Aunts whom I am truly grateful for having in my life.

Acknowledgement

First and foremost, I have to thank my parents for their love and support throughout my life. Thank you both for giving me strength to reach for the stars and chase my dreams. My brothers, aunties, uncles and cousins deserve my wholehearted thanks as well.

I would like to sincerely thank my thesis advisers, Dr Ali Ahmed and Dr Faisal Mohammed, for their guidance and support throughout this study and especially for their confidence in me. To all my friends, thank you for your understanding and encouragement in many, many moments of crisis. (our friendship makes my life a wonderful experience. I cannot list all the names here, but you are always on my mind.

Thank you, Allah, for always being there for me.

Abstract

Cloud computing is a recent technology that uses the Internet, central servers to organize the data and applications, which the user can access, cloud computing enables individuals and organizations to gain access to huge computing resources without capital investment. one of the main drawbacks of cloud computing is data security. when data migrate to the cloud, is fully controlled by cloud service provider not by the data owner. As a result, user data is not secure at the server side. Security goals of data cover three points namely: Availability, Confidentiality, and Integrity [1]. This thesis handled Confidentiality of file using hybrid of symmetric encryption algorithms (AES with key (128), AES with key (256) and Blowfish),the proposed technique based on providing data confidentiality by the software, each user wants to upload to, or download a file from the cloud storage which must be passed to the software. Where the software performs encoding process by splitting it into three blocks with different sizes and selecting algorithms randomly to encrypt blocks of file before uploading and saving the information's of encoding process in local database, also the software performs the decoding process by retrieving the information's of encoding process to decrypt the file after downloading it from the cloud storage.the thesis was checked the randomness of sequence binary bits of file, and evaluate the time of the encoding and decoding processes of this proposed method and compared it with time of encryption algorithms: AES-128, AES-256, Blowfish.

المستخلص

الحوسبة السحابية هي تقنية حديثة تستخدم الإنترنت ، وخوادم مركزية لتنظيم البيانات والتطبيقات ، والتي يمكن للمستخدم الوصول إليها ، والحوسبة السحابية تمكن الأفراد والمؤسسات من الوصول إلى موارد ضخمة دون استثمار رأس المال. واحدة من العوائق الرئيسية للحوسبة السحابية هي أمن البيانات. فعندما يتم ترحيل البيانات إلى التخزين السحابي ، يتم التحكم فيها بالكامل من قِبل مقدم الخدمة وليس بواسطة مالك البيانات. ونتيجة لذلك ، تكون بيانات المستخدم غير آمنة على جانب الخادم. تغطي الأهداف الأمنية للبيانات ثلاث نقاط هي :اتاحة البيانات ،السرية ، و سلامة البيانات. تناولت هذه الرسالة سرية الملف باستخدام هجين من خوارزميات التشفير المتماثلة(AES-256, AES-128 and Blowfish)، التقنية المقترحة تقوم بتوفير سرية البيانات من قبل البرنامج. فإذا اراد مستخدم رفع ملف ، أو تنزيله من التخزين السحابي فيقوم اولا بتمريره إلى البرنامج.حيث يقوم البرنامج بتنفيذ عملية التشفير و يتم فيها تقسيمه الى ثلاثة كتل مختلفة الاحجام و من ثم اختيار الخوارزميات بصورة عشوائية لتشفير هذه الكتل و حفظ معومات عملية التشفير في قاعدة بيانات قبل رفع الملف، و ايضا يقوم البرنامج بعملية فك التشفير باسترجاع معلومات عملية التشفير لفك تشفير الملف بعد تنزيله من التخزين السحابي. هذه الدراسة قامت باختبار عشوائية الملف كما قامت بتقييم زمن عملية التشفير و فك التشفير و مقارنته مع زمن خوارزميات التشفير:.AES-128, AES-256, Blowfish

Table of Contents:

Table of Figures:

List of Tables:

Abbreviation:

Notations	Description
AES	Advance Encryption Standard
B1	The First Block
B2	The Second Block
B3	The Third Block
Bid	File Block id
Bsize	Block size
D1	The First Decryption Algorithm
D2	The Second Decryption Algorithm
D3	The Third Decryption Algorithm
DES	Data Encryption Standard
E1	The First Encryption Algorithm
E2	The Second Encryption Algorithm
E3	The Third Encryption Algorithm
Eid	Encryption Algorithm id
Fid	File id
K128	Key of AES-128
K256	Key of AES-256
Kalg	Key of Algorithm
KB	Keys of Blowfish

Chapter One

Introduction

1.1 Background:

Cloud computing is the use of computing resources that is delivered as a service over a network. NIST defines, “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” [2]. Cloud computing comes in three categories such as Software as a Service (SaaS), Infrastructure as a service (IaaS), Platform as a Service (PaaS).The SaaS provides application software which the user can use. The PaaS provides the platform for the user to do his operation. The IaaS provide physical or virtual devices for user. And each provides different services to the user. The cloud is available in four-deployment model namely [1] (Public Cloud, Private Cloud, Community Cloud, and Hybrid Cloud).

Cloud delivers storage as a service (STaaS) all over the internet. STaaS is a service where data is remotely maintained, managed, and backed up [2]. Cloud storage allows user to access broad range of application and resources immediately, which are hosted by others [2].

Data in the public storage cloud may be expose to disclosure by unauthorized users, Confidentiality of data is a very important aspect of information security. Confidentiality of data is to be ensured by cryptography technique [2]. Cryptography is a technique which is intended to transform the data and can be used to provide various security related concepts such as confidentiality, data integrity, etc. [3]

As the central data storage is the key facility of the cloud computing it is of prominent importance to provide the security. The art and science of concealing the messages to introduce secrecy in information security is recognized as cryptography. Security goals of data cover three points namely: Availability, Confidentiality, and Integrity. Cryptography, in modern days is considered grouping of three types of algorithms. They are

1) Symmetric-key algorithms
2) Asymmetric-key algorithms
3) Hash functions

Symmetric algorithms use the same key for encryption and decryption. This is termed as secret key. With the same key messages are encrypted by the sender and decrypted by the receiver. It contains algorithms like Data Encryption Standard (DES), Advanced Encryption Standard (AES), Ron's Code (RCn), and Triple DES, Blowfish etc.

Asymmetric algorithms use different keys. One key (public) is used for encryption and other (private key) is used for decryption. This is named as public key. Public key is known to public and private key is known to the user. It comprises various algorithms like Rivest, Shamir, &Adleman (RSA), Digital

Signature Algorithm (DSA), Elliptic Curve (EC), Diffi-Hillman (DH), El Gamal etc.
The Hash functions use a mathematical transformation to irreversibly "encrypt" information. It contains algorithms like Message Digest, Secure Hash Algorithm.
AES is most frequently used encryption algorithm today this algorithm is based on several substitutions, permutations and linear transformations, each executed on data blocks of 16 byte. As of today, no practicable attack against AES exists. Therefore, AES remains the preferred encryption standard for governments, banks and high security systems around the world [4].
Many researchers in this field have used the technique of splitting file into slices and make sequence number for each slice and then encrypting these slices separately using one or more cryptography algorithms with different keys. This technique has recently become popular for data storage security in cloud storage [1].

1.2 Problem Statement:

Today, the file that stored in public cloud storage is exposing to disclosure from unauthorizedaccess,the confidentiality of this file is done by a cryptography technique. One encryption algorithm easy to be attack, and when file migrate to the cloud is fully controlled by cloud service provider not by the data owner, somewhen the provider side of cloud services exposed to attack.

1.3 Research Importance:

Cloud computing is a technology, which provides low cost, scalable computation capacity and services to enterprises on demand for expansion. However, the sharing of the resources in an open environment which leads to the security problems [examples: data integrity, data leaking, etc.] [5], all of these problems make user to reluctance to use cloud computing facility.Therefore, it is very important to protect this data and never expose it to leaking or modification.

1.4 Research Hypotheses:

1- Using hybridencryption algorithmto encrypt the file before uploading processwill provide user privacy in high degree.

2- the thesis approaches are insufficient to protect data

1.5 Research Objectives:

- Providing confidentiality of file that stored in public cloud storage with trusted level of security.
- Enhance the prevention of files against cryptanalytic attacks with high level
- Enhance the prevention of files against brute force with high level

1.6 Research Scope:

This thesis applied to provide file confidentialityto the user,when user wants to store the file in public cloud storage using local PC.

1.7 Research Methodology:

A big issue when adopting the cloud, which is security of data. The cloud has manychallenges and issues, but among that security data is the main concern. Mostly,security goals of data cover three points namely: Availability, Confidentiality, and Integrity. This methodology highlighted the confidentiality of data, cryptography techniquesarepreferred for data confidentiality and most of the techniques are outdated [2]. Therefore, thisthesis proposed a techniquethat ensures data confidentiality,and prevent data leakage as well.

The technique based on providing data confidentiality by software. Each user wants to upload to,or download a filefromthe cloud storagewhich must be passed to the software. The software performs Encoding processbefore uploading and Decoding process afterdownloading file from the cloud storage.

1.8Thesis Layout:

- **Chapter one:**gives introduction about the thesis, defining the problem, objectives, methodology and scope.
- **Chapter two:**contains general background about cloud computing. and the related studies and techniques that used in confidentiality of data.
- **Chapter three:**contains the methodology which explain the proposed technique.
- **Chapter four**: contains the implementation of technique that will be used.
- **Chapter five:** contains the results and recommendations.

Chapter Two
Literature Review

2.1 Cloud Computing:

Cloud Computing is a group of computers and servers linked together over the internet to maintain data and applications. It refers to manipulating, designing and accessing the applications online. It allows consumers and businesses to use applications without installation and access their personal file from any computer with the help of internet. It is architecture for providing computing services via internet on demand and pay per use access to a pool of shared resources for the network storage, services and applications. It is totally an internet-based technology in which client data is stored and maintained in data center of cloud provider like Google, Amazon, and Salesforce.com etc. The client can access your application from anywhere. The cloud computing can be seen as the important change of information industry and will make more impact on the development of information technology for the society [5].

2.2 Cloud Deployment Models:

2.2.1 Public Cloud

In a public cloud, IT resources are made available to the public organizations and are owned by the Cloud service provider. The cloud services are made accessible to everyone via standard internet connection. In a public cloud, a service provider makes IT resources such as applications, storage capacities available to any consumer. This model is considered as an on-demand and pay-per use environment, where there are no on-site infrastructure or management requirements. These benefits come with certain risks such as no control over the resources, data security, network performance and interoperability.

2.2.2 Private Cloud

In a private cloud, the cloud infrastructure operates separately for each organization and is not shared with any other organizations. This cloud model offers the greatest level of security and control. The two variations are as follows, o on premise private cloud: This is also known as internal clouds and are hosted by an organization within their own data centers. This model provides a more standardized process, but is limited in terms of size and scalability. This is best suited for applications which require complete control of the infrastructure and security. Externally-hosted private cloud: This type of private cloud is hosted externally with a cloud provider, where the provider facilitates an exclusive cloud environment for a specific organization with full guarantee of privacy or confidentiality.

2.2.3 Community Cloud.

A community cloud is a multi-tenant infrastructure that is shared among several organizations from a specific group with common computing concerns. Such concerns might be related to regulatory compliance, such as audit requirements, or may be related to performance requirements, such as hosting applications that require a quick response time. The goal of a community cloud

is to have participating organizations realize the benefits of a public cloud but with the added level of privacy, security and policy compliance usually associated with a private cloud. The community cloud can be either on-premises or off-premises, and can be governed by the participating organizations or by a third-party managed service provider.

2.2.4 Hybrid Cloud

In a hybrid cloud environment, the organization consumes resources from both private and public clouds. For the maintenance of service levels, the public cloud resources are imbibed with the private cloud resources. Organizations use their computing resources on a private cloud for normal usage, but access the public cloud for peak load/high requirements. This ensures that a sudden increase in computing requirement is handled gracefully [6].

2.3 Cloud Service Models:

2.3.1 Infrastructure as a Service (IaaS):

IaaS this model provides basic storage and computing capabilities as standardized services over the network. Servers, storage systems, networking equipment, data center space etc. are pooled and made available to handle workloads. The capability provided to the customer is to rent processing, storage, networks, and other fundamental computing resources where the customer is able to deploy and run arbitrary software, which can include operating systems and applications. The customer does not manage or control the underlying cloud infrastructure but has the control over operating systems, storage, deployed applications, and possibly select networking components (e.g., firewalls, load balancers etc.). Some examples of IaaS are: Amazon, GoGrid, 3 Tera etc [7].

2.3.2 Platform as a Service (PaaS):

PaaS in this model, a layer of software or development environment is encapsulated and offered as a service, upon which other higher levels of service are built. The customer has the freedom to build his own applications, which run on the provider's infrastructure. Hence, capability is provided to the customer to deploy onto the cloud infrastructure customer-created applications using programming languages and tools supported by the provider (e.g., Java, Python, .Netetc.). Although the customer does not manage or control the underlying cloud infrastructure, network, servers, operating systems, or storage, but he/she has the control over the deployed applications and possibly over the application hosting environment configurations. To meet manageability and scalability requirements of the applications, PaaS providers offer a predefined combination of operating systems and application servers, such as LAMP (Linux, Apache, MySql and PHP) platform, restricted J2EE, Ruby etc. Some examples of PaaS are: Google's App Engine, Force.com, etc.

2.3.3 Software as a Service (SaaS):

SaaS the capability provided to the consumer is to use the provider's applications running on a cloud infrastructure and accessible from various client devices through a thin client interface such as web browser. In other words, in this model, a complete application is offered to the customer as a

service on demand. A single instance of the service runs on the cloud and multiple end users are services. On the customers' side, there is no need for upfront investment in servers or software licenses, while for the provider, the costs are lowered, since only a single application needs to be hosted and maintained. In summary, in this model, the customers do not manage or control the underlying cloud infrastructure, network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings. Currently, SaaS is offered by companies such as Google, Salesforce, Microsoft, Zohoetc [7].

2.4 Public Cloud Storage:

Cloud Storage provides cost-effective services to individual users as well as organization. It provides huge amount of space to outsource the data to the cloud. Organization and enterprises do not possess full infrastructure to maintain their data with their premises. Data outsourcing helps to effectively maintain their data in cloud storage [2]. The following examples for cloud storage:

2.4.1 Amazon S3:

Amazon Simple Storage Service (Amazon S3) is an object storage service that offers industry-leading scalability, data availability, security, and performance. This means customers of all sizes and industries can use it to store and protect any amount of data for a range of use cases, such as websites, mobile applications, backup and restore, archive, enterprise applications, IoT devices, and big data analytics. Amazon S3 provides easy-to-use management features so you can organize your data and configure finely-tuned access controls to meet your specific business, organizational, and compliance requirements. Amazon S3 is designed for 99.999999999% (11 9's) of durability, and stores data for millions of applications for companies all around the world [7].

2.4.2 Google Drive:

Google Drive is a free cloud-based storage service that enables users to store and access files online, and is considered a file storage and synchronization service developed by Google. Launched on April 24, 2012, Google Drive allows users to store files on their servers, synchronize files across devices, and share files. In addition to a website, Google Drive offers apps with offline capabilities for Windows and macOS computers, and Android and iOS smartphones and tablets. Google Drive encompasses Google Docs, Sheets and Slides, an office suite that permits collaborative editing of documents, spreadsheets, presentations, drawings, forms, and more. Files created and edited through the office suite are saved in Google Drive [8].

2.5 Cloud Computing Security Issues:

There are numerous security issues for cloud computing as it encompasses many technologies including networks, databases, operating systems, virtualization, resource scheduling, transaction management, load

balancing, concurrency control and memory management. Therefore, security issues for many of these systems and technologies are applicable to cloud computing.

The threats to information assets residing in the cloud can vary according to the cloud delivery models used by cloud user organizations. There are several types of security threats to which cloud computing is vulnerable. Table 2.1 provides an overview of the threats for cloud customers categorized according to the confidentiality, integrity and availability (CIA) security model and their relevance to each of the cloud service delivery model [9].

Threat	**Description**
Confidentiality	
Insider user threats: • Malicious cloud provider user • Malicious cloud customer user • Malicious third-party user (Supporting either the cloud provider or customer organizations)	The threat of insiders accessing customer data held within the cloud is greater as each of the delivery models can introduce the need for multiple internal users: SaaS – cloud customer and provider administrator's PaaS- application developers and test environment managers IaaS- third party platform consultants
External attacker threats: • Remote software attack of cloud infrastructure • Remote software attack of cloud applications • Remote hardware attack against the cloud • Remote software and hardware attack against cloud user organizations' endpoint software and hardware • Social engineering of cloud provider users, and cloud customer users.	The threat from external attackers may be perceived to apply more to public Internet facing clouds, however all types of cloud delivery models are affected by external attackers, particularly in private clouds where user endpoints can be targeted. Cloud providers with large data stores holding credit card details, personal information and sensitive government or intellectual property, will be subjected to attacks from groups, with significant resources, attempting to retrieve data. This includes the threat of hardware attack, social engineering and supply chain attacks by dedicated attackers.
Data leakage: ❖ Failure of security access rights across multiple domains ❖ Failure of electronic and physical transport systems for cloud data and backups.	A threat from widespread data leakage amongst many, potentially competitor organizations, using the same cloud provider could be caused by human error or faulty hardware that will lead to information compromise.

Integrity	
Data segregation: • Incorrectly defined security perimeters. • Incorrect configuration of virtual machines and hypervisors.	The integrity of data within complex cloud hosting environments such as SaaS configured to share computing resource amongst customers could provide a threat against data integrity if system resources are effectively segregated.
User access: • Poor identity and access management procedures.	Implementation of poor access control procedures creates many threat opportunities, for example that disgruntled ex-employees of cloud provider organizations maintain remote access to administer customer cloud services, and can cause intentional damage to their data sources.
Data quality: • Introduction of faulty application or infrastructure components	The threat of impact of data quality is increased as cloud providers host many customers' data. The introduction of a faulty or misconfigured component required by another cloud user could potentially impact the integrity of data for other cloud users sharing infrastructure.
Availability	
Change management: • Customer penetration testing impacting other cloud customers. • • Infrastructure changes upon cloud provider, customer and third party systems impacting cloud customers.	As the cloud provider has increasing responsibility or change management within all cloud delivery models, there is a threat that changes could introduce negative effects. These could be caused by software or hardware changes to existing cloud services.
Denial of service threat: • Network bandwidth distributed denial of service • Network DNS denial of service • Application and data denial of service**.**	The threat of denial of service against available cloud computing resource is generally an external threat against public cloud services. However, the threat can impact all cloud service models as external and internal threat agents could introduce application or hardware components that cause a denial of service.

Table 2-1: A list of cloud security threats [9]

2.6 Security Techniques:

Recently, many researchers in this field have been interested in protecting data stored in outsourced, they used many techniques. Some of them used data encryption and some of them used data Obfuscation.

2.7 Cryptography:

Cryptography is an effective tool that helps to protect the data from unauthorized access while data at rest in cloud server. Cryptography is a method of storing and transmitting data in a particular form so that only those for whom it is intended can read and process it. It is the process of encryption and decryption. Cryptographic techniques are classified into Conventional and Public key cryptography. Conventional cryptography is also referred as symmetric key cryptography. The same key is used for encryption and decryption in symmetric key cryptography. Public key cryptography is called as asymmetric key cryptography. Public key and private key are used for encryption and decryption respectively [2].

2.7.1 Symmetric Encryption Algorithm:

In symmetric Cryptography the key used for encryption is similar to the key used in decryption. Thus, the key distribution has to be made prior to the transmission of information. The key plays a very important role in symmetric cryptography since their security directly depends on the nature of key i.e. the key lengthiest. There are various symmetric key algorithms such as DES, TRIPLE DES, AES, RC4, RC6 and BLOWFISH.

2.7.2 Asymmetric Encryption Algorithm:

In Asymmetric Key encryption, two different keys are used for encryption and decryption- Public and Private. The public key is meant for general use so it is available to anyone on the network. Anyone who wants to encrypt the plaintext should know the Public Key of receiver. Only the authorized person can be able to decrypt the cipher text through his own private key. Private Key is kept secret from the outside world. There are various asymmetric key techniques such as RSA, Diffie-Hellman Key exchange and elliptic curve.

For example, A wants to send message to B. The following steps are involved
A and B should know public key of each other but private keys are kept secret.

a) A encrypts a Plain Text message for B by using B's public key.
b) A transmits the encrypted message (Cipher Text) to B.
c) B receives the cipher text and decrypts it using its own private key.
d) B gets the Plain Text message [10].

2.7.3 A Comprehensive Evaluation of Cryptographic Algorithms: DES, 3DES, AES, RSA and Blowfish [11]:

A study conducted in analyzing algorithms performance corresponding the file size, shows that the performance of blowfish takes least time for

encryption and decryption, AES shows relatively moderate time for encryption and decryption no matter the size of file, and shows that RSA takes highest time for encryption and decryption as shown in figure 2.1:

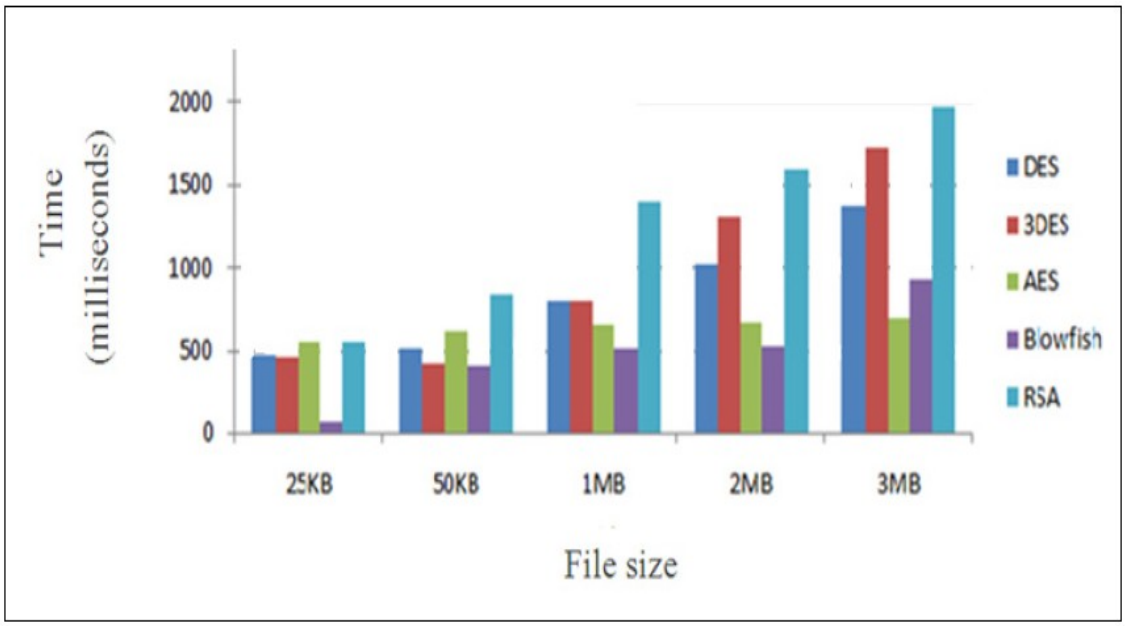


Figure 2.1 a Comprehensive Evaluation of Cryptographic Algorithms [11]

2.8 Related Works:

2.8.1 Data Encryption Techniques:

Definition: Encryption is the method by which plaintext or any other type of data is converted from a readable form to an encoded version that can only be decoded by another entity if they have access to decryption key [12].

A Secured Storage using AES Algorithm and Role Based Access in Cloud [13], The main contribution of this paper designed a prototype which uses a encryption technique (AES algorithm) to store data and Retrieve data using access control (Role Based Access).In this study, instead of applying encryption technique to the whole dataset, they were applying attribute-based encryption to only any confidential data in the database and the same encrypted data sent to cloud server. Accessing of data by authorized users via access control methods and polices.

The proposed scheme provided two-factor protection mechanism to enhance the confidentiality of outsourced data. If a user wants to recover the outsourced data, this user is required to hold sufficient attribute secret keys with respect to the access policy and authorization key with regard to the outsourced data. In addition, the proposed scheme provided the user-level revocation for data owner in attribute-based data access control systems.

Enhancement of Cloud Computing Security with Secure Data Storage using AES [4], This paper presented the symmetric cryptographic algorithm named as AES (Advanced Encryption Standard) to protect data

in cloud database server.

2.8.2 Data Obfuscation Technique:

Definition: Obfuscation is the process of hiding the original value of data [2].

Confidentiality Technique for Enhancing Data Security in Public Cloud Storage using Data Obfuscation [Confidentiality- Obfuscation], this paper proposed a confidentiality technique named as WMRADO (Word Magical Rolling Alpha Digits Obfuscation) to enhance the security of data in cloud. This technique is applied in plaintext in both ways such as word by word and line by line obfuscation. WMRADO technique is based on word by word obfuscation. WMRADO Technique improved classical obfuscation techniques by integrating substitution, transposition and ASCII values. This technique is used to hide the original value of data and should not be reversible. Initially the Word (W) lists are generated from the plain text. In the corresponding lines, characters(C) are converted to ASCII value and that ASCII (ASC) value is multiplied with that character position where it appeared in the word. After that, then multiplied value of individual character is added with another character value and it produces a Numerical Code (NC) to the corresponding word. The look-up table (LT) is generated. It maintains the line and the corresponding NC value which is derived from the lines. The Modulo Operation (MO) is performed on NC by 64 to get the remainder and quotient value.

The proposed method enhanced the data security and also reduces the cost of data storage while uploading the data to the cloud storage compared to the MONcrypt technique (calculated by using Google Cloud storage cost).

2.8.3 Splitting, Merging and Hybrid Cryptography Techniques

Security Enhancement for Data Migration in the Cloud [14], in this paper proposed a model that enhanced data security and privacy by combining Advanced Encryption Standard-256 (AES), Information Dispersal Algorithms (IDA) and Secure Hash Algorithm-512.

The encoding operation includes: AES-256 Encryption to ensure data confidentiality, IDA with Cauchy Reed–Solomon code to break the encrypted data into n slices such that we can recover from m and then SHA-512 Hashing algorithm for signature. However, the decoding operation includes a verification process to check data slice integrity, IDAs to reconstruct the encrypted data from m slices, and, finally, the decryption process to recover the original data.

Compared to other approaches to secure data privacy, their algorithm achieved far a greater degree of security and also better performance for small and large data files.

Secure file storage in cloud computing using hybrid cryptography algorithm [15], in this proposed method the file being encrypted will be

sliced into n slices. Each of the file slices is encrypted using Blowfish key provided by the user for each slice. The key will be encrypted using SRNN public key after encryption, at the decryption phase the client will give n SRNN private keys, as indicated by the quantity of cuts (n) made amid the encryption stage. Blowfish key is decoded at the server end utilizing the SRNN private key particular to the cut. Utilizing the relating unscrambled Blowfish keys, document cuts put away at server are decoded. The unscrambled cuts will be converged to produce unique record.

Enhancement of cloud computing security in health care sector [6], the proposed system provided two-layer protection to secure the Electronic Health Records. In the first layer, the images and the text files are encrypted using Advanced Encryption Standard (AES). In the second layer, the encrypted files are divided into n files. These n files are then stored in the cloud. The original Electronic Health Record can be decrypted only if the n files are merged. For splitting and merging the cipher texts a sequence key will be used.

2.8.4 Other Methods:

Secure Data Storage over Cloud Using Content Features Processing [16], the proposed project provided effective security and integrity by using Shamir's secret hashing algorithm and Elliptic Curve Cryptography (ECC) algorithm. It also provides data verifier scheme which assure data integrity not only at data user side but also assure data admin too about confidentiality and data security on a cloud server. Data admin upload their data on a cloud and generate a hash value on original text file using SHA-1 algorithm, admin can download the text file from a cloud and applied hashing algorithm on it again to check the data integrity, if hash value get match means no data get tampered confidentiality is maintain on cloud server. Else if data admin want can directly upload text file on cloud without integrity checking. Data Verifier scheme used at data admin level which provide the data security on a cloud as well as assure admin about its uploaded data not get manipulated by attacker. After uploading the text file on a cloud server, Elliptic Curve Cryptography (ECC) algorithm used to encrypt the text file using secrete key encryption. The key mail to a respective user whom data admin wants to share text file.

Data user can view a share files by data admin, and can download a share file too through secrete decryption key present on a data user given mail Data user can again check the integrity of original download text file content by using SHA-1 hash value algorithm. This scheme is known as data verifier which match the original uploaded text file hash value with downloaded text files hash value. Both hash value get match means no data get manipulated.

Data user can also search a file on a cloud server using a search term keyword which gives relevant files list to a data user which are not shared by data admin.

Providing Confidentiality and Integrity on Data Stored in Cloud Storage by Hash and Meta-data Approach [2], in this paper Authors were proposed a new approach for securely storing our data in cloud and integrity checking mechanism. To store the local file securely in the cloud, firstly encrypt the local file using AES-256 encryption algorithm and create the meta-data of that file, after that files are uploaded in the cloud. To provide integrity checking they generated the hash of the local file using SHA256 hashing algorithm.

2.9 Summary of Related Works:

Paper name	Author	Technique	Result
A Secured Storage using AES Algorithm and Role Based Access in Cloud, 2017	M.Saraswathi,T.Bhuvaneswari	AES, Role Based Access	Proposed scheme enhanced the confidentiality of outsourced data.
Enhancement of cloud computing security with secure data storage using AES, 2016	Vishal R.Panchli et al	AES Algorithm	AES encryption algorithm had minimal storage space and high performance without any weaknesses and limitations while other symmetric algorithms have some weaknesses and differences in performance and storage space. And required memory for AES algorithm is less than the Blowfish.
Confidentiality Technique for Enhancing Data Security in Public Cloud Storage using Data Obfuscation, 2016	D.I.George Amalarethinam et al	MRADO Technique	MRADO technique enhanced the data security and also reduced the cost of data storage compare to the existing MONcrypt technique
Security Enhancement for Data Migration in the Cloud, 2017	Lin You et al	AES-256, Hash-512	The maximum encoding time is 14.15 s for a large data file equal to 347,778 KB with (200, 254) threshold. Their observed for different (m, n) configurations and data less than 347,778

			KB that our encoding algorithm yields the best performance, since the average encoding time is 1.966 s.
Secure file storage in cloud computing using hybrid cryptography algorithm, 2017	B.Swathi et al	SRNN, Blowfish algorithm	Split file and encrypt the key of blowfish made file more secure.
enhancement of cloud computing security in health care sector, 2017	R. Anbuselvi et al	AES, Split-merge File	Solution it had speed and computational efficiency to handle encryption of large volumes of data. In AES, the longer the key length, the stronger the encryption.
Providing Confidentiality and Integrity on Data Stored in Cloud Storage by Hash and Meta-data Approach, 2017	Jeet vyaset al	Hash and Meta-data Approach with AES	Enhanced the security of the data which is stored in cloud. As, we store the encrypted file in cloud and also provide integrity checking mechanism.

Table 2-2: Summary of Related Works

2.10 Comparative Studies and Open Issues:

- "Enhancement of Cloud Computing Security with Secure Data Storage using AES" [4], this paper used AES encryption algorithm to protect data in cloud database server, AES has speed and computational efficiency to handle encryption of large volumes of data. Encryption and decryptions processes were done in cloud server, which is means data is not protected when it migrates from user to cloud server.
- "Enhancement of cloud computing security in health care sector" [6], this paper provided two-layer protection, in the first layer, the images and the text files are encrypted using AES. In the second layer, the encrypted files are divided into n files using sequence key (greater than 8 bits), these n files are then stored in the cloud, and this makes retrieving process of file from the server it is very difficult and complex.
- "Providing Confidentiality and Integrity on Data Stored in Cloud Storage by Hash and Meta-data Approach" [2], in this paper encrypt the local file using AES-256 and create the meta-data of that file, after that files are uploaded in the cloud. To provide integrity checking they generated the hash of the local file using SHA256 hashing algorithm, all these file (encrypted file, hashed file and meta-data) are stored in cloud, these generated files will take up a large storage space.

Chapter Three

Research Methodology

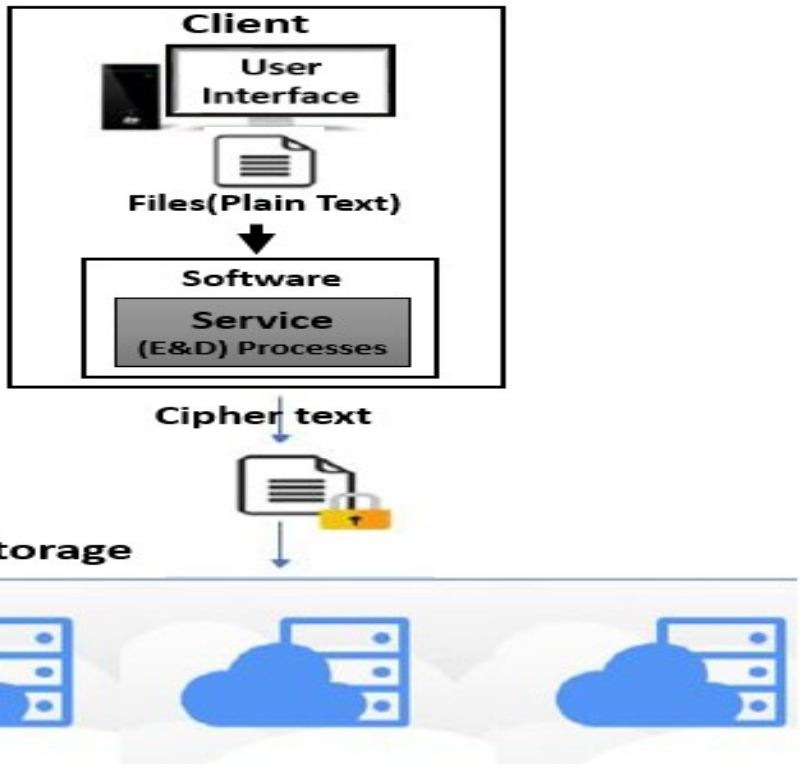
3.1 Introduction:

This chapter shows the methodology that providing confidentiality of File storage in public cloud using symmetric encryption algorithms (AES-128, AES-256, and Blowfish).

3.2 Research Methodology:

A big issue when adopting the cloud, which is security of data. The cloud has many challenges and issues, but among that security data is the main concern. Mostly, security goals of data cover three points namely: Availability, Confidentiality, and Integrity. This methodology highlighted the confidentiality of data, cryptography techniques are preferred for data confidentiality and most of the techniques are outdated [1]. Therefore, this thesis proposed a technique that ensures data confidentiality, and prevent data leakage as well.

The technique based on providing data confidentiality by software. Each user wants to upload to, or download a file from the cloud storage which must be passed to the software. The software performs Encoding process before uploading and decoding process after downloading file from the cloud storage, as shown in figure 3-2:



-
:

Figure 3 2 Research Methodology in General

3.3 Proposed Technique Components:

3.3.1 User Interface:

The interface enables the user to select specific file to be encrypt and upload or download the file and decrypt it.

3.3.2 The Software:

The software is designed to provide service contain encoding and decoding (E&D) processes. The encoding and decoding processes used hybrid symmetric encryption algorithm to provide confidentiality of file.

:Encoding Process 3.3.2.1

The encoding process consists of seven steps: reading bytes, splitting file, choosing algorithm, key generation, encryption, saving data encoding in local database, merge file (write bytes).

During the encoding process firstly reads bytes from the input file after that splits it into three blocks with different sizes randomly, then one of the three algorithms are chosen randomly to encrypt each block, after that keys generated for each algorithm (AES (128), AES (256) and Blowfish), and then saves the information of encoding process into local database, finally the blocks merged into one file to upload to cloud storage. as shown in figure 3-3:

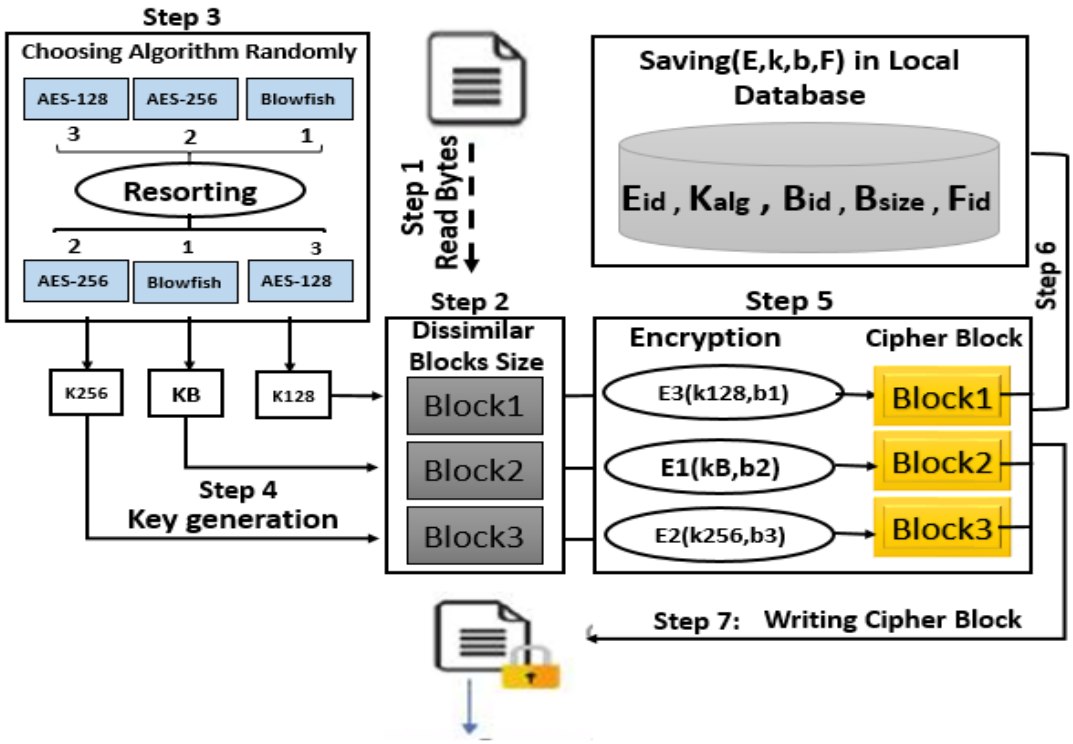


Figure 3-3: Encoding Process

:Steps of Encoding Process 3.3.2.1.1

Step 1: Reading Bytes

In this Step the software reads bytes form the input file continuously.

Step 2: Splitting File

In this step three outsources produced to be used in writing the bytes of blocks. Before writing bytes three numbers must be generated randomly that represent the sizes of blocks (size1, size2 and size3). Then write byte in the out sources depending on the sizes of blocks, which read in the step 1.

After writing and reading processes, three blocks are produced with different sizes (Block1, Block2 and Block3).

Step 3: Choosing Algorithm

There are symmetric algorithms (AES (128), AES (256) and Blowfish), each algorithm has a specific identifier number, these algorithms are rearranged every time the client upload new file on cloud storage server.

Step 4: Key Generation

In this step the software generates three different keys for each algorithm:

- AES (128) key with 128 bits in size.
- AES (256) key with 256 bits in size.
- Blowfish key with 64 bits in size.

Step 5: Encryption

After keys generation step, both keys and algorithms are used to encrypt each of blocks (block1, block2 and block3) randomly.

When encrypting blocks of file is finished that lead to produce cipher blocks.

Step 6: Saving Information of Encoding Process

In this step the software saves all of the information about encoding process in local database that will be needed in decoding process.

The information of encoding process is:

- File id
- Block id
- Blocks sizes
- Algorithms of encryption
- Keys of algorithms

Step 7: Merging (writing bytes)

In this step the software writes cipher blocks into one source. When writing is done encrypted file will be produced.

After completing encoding process the software will upload confidential file to public cloud storage.

3.3.2.1.2 Encoding Process Algorithm:

Output: encrypted file

Input: bytes of file

Generate random size (s1, s2, s3)

//begin while loop

While (read_byte [] != -1)

If (read <= s1)

Block1 ← block1+read

Else if (read > s1 and read <=s2)

Block2 ← block2+read

Else if (read > s2)

Block3 ← block3+read

//end while loop

Resorting Algorithms indexes ()

Generating keys ()

Cipher Block1 ← Algorithm1 (key1, block1)

Cipher Block2 ← Algorithm2 (key2, block2)

Cipher Block3 ← Algorithm3 (key3, block3)

File ←Merge (Cipher Block1, Cipher Block2, Cipher Block3)

Save (key1, key2, key3, s1, s2, s3, Algorithm1, Algorithm2, Algorithm3,file id)

Return file

3.8.2.2Decoding process:

The decoding process consists of six steps: reading bytes, getting blocks sizes, re splitting file, getting algorithms and keys, decryption, merging file (write bytes).

In the beginning of the decoding process, bytes are read from the input file (downloaded file), and getting blocks sizes from local database to re spilt it into three blocks with the same sizes in encoding process, and then the software will gets algorithm and key of each algorithm to decrypt each block by these algorithms (AES-128, AES-256, Blowfish) depending on algorithm ID, finally the three blocks will be merged into one file (writing bytes). As shown in figure 3-4.

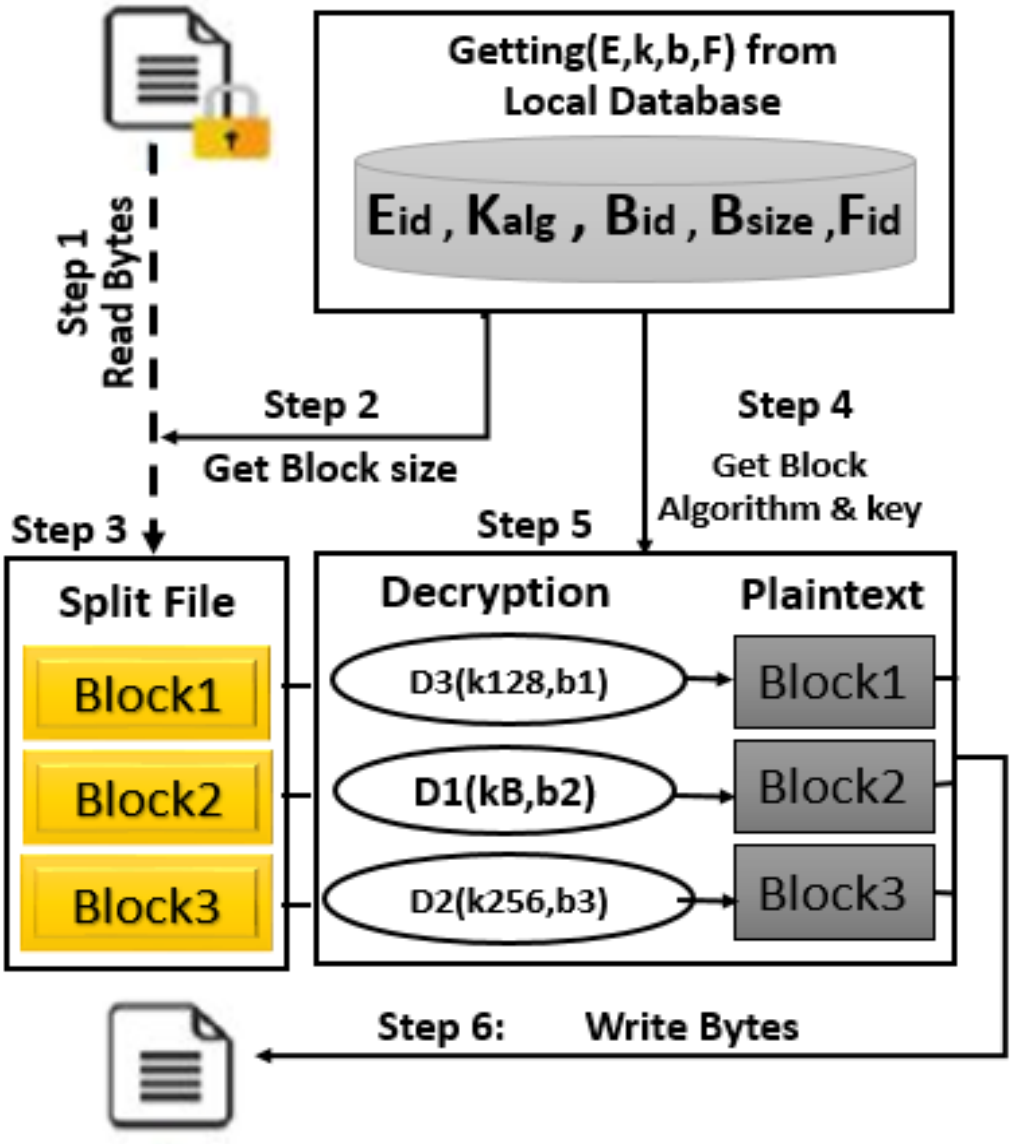


Figure 3-4: Decoding Process

:Steps of Decoding Process 3.8.2.2.1

Step 1: Reading Bytes

In this step the software reads bytes form the input file continuously, each time only one byte will be read.

Step 2: Getting Blocks Sizes

In this step the software gets blocks sizes form the local database to re spilt encrypted file into three blocks with the same sizes in encoding process.

Step 3: Splitting File

Three outsources produced in this step by software that will be used for writing the bytes of cipher blocks.

After getting blocks sizes software will write bytes in outsources depending on these sizes, the reading and writing are sequentially performed, and when writing and reading processes are finished, they produce three cipher blocks (Block1, Block2, and Block3).

Step 4: Get Block Algorithm & Key

In this step the software retrieves algorithm that was used in encryption and also will retrieve the key of encryption algorithm from local database. The algorithm and key are used as input for the next step (Decryption).

Step 5: Decryption

In this step the software will use algorithm & key to decrypt cipher blocks of the file, each cipher block will be decrypt with corresponding algorithm & key. When decryption is finished decrypted blocks will be produced.

Step 6: Merging (writing bytes)

In this step the software writes decrypted blocks into one outsource. When writing is finished decrypted file will be produced.

3.8.2.2.2 Decoding Process Algorithm:

Output: decrypted file

Input: encrypted file

Get size (s1, s2, s3)

//begin while loop

While (read byte []! = -1)

If (read <= s1)

Cipher Block1 cipher block1+read

Else if (read > s1 and read <=s2)

```
Cipher Block2 cipher block2+read
        Else if (read > s2)
Cipher Block3 cipher block3+read
    //end while loop
Get algorithms ()
Get keys ()
    Block1  Algorithm1 (key1, cipher block1)
    Block2  Algorithm2 (key2, cipher block2)
    Block3  Algorithm3 (key3, cipher block3)
File Merge (Block1, Block2, Block3)
Return file
```

Chapter Four

Implementation

4.1 Introduction:

This chapter shows the tools that used in this thesis, and also clarify tools integration by chart.

4.2 Tools of Encryption Processes:

4.2.1 AES Encryption Algorithm:

The Advanced Encryption Standard (AES) was published by NIST (National Institute of Standards and Technology) in 2001. AES is a symmetric block cipher that is intended to replace DES as the approved standard for a wide range of applications. NIST selected Rijndael as the proposed AES algorithm. The two researchers who developed and submitted Rijndael for the AES are both cryptographers from Belgium: Dr. Joan Daemen and Dr.VincentRijmen.

4.2.2.1 NIST's Requirements for the AES:

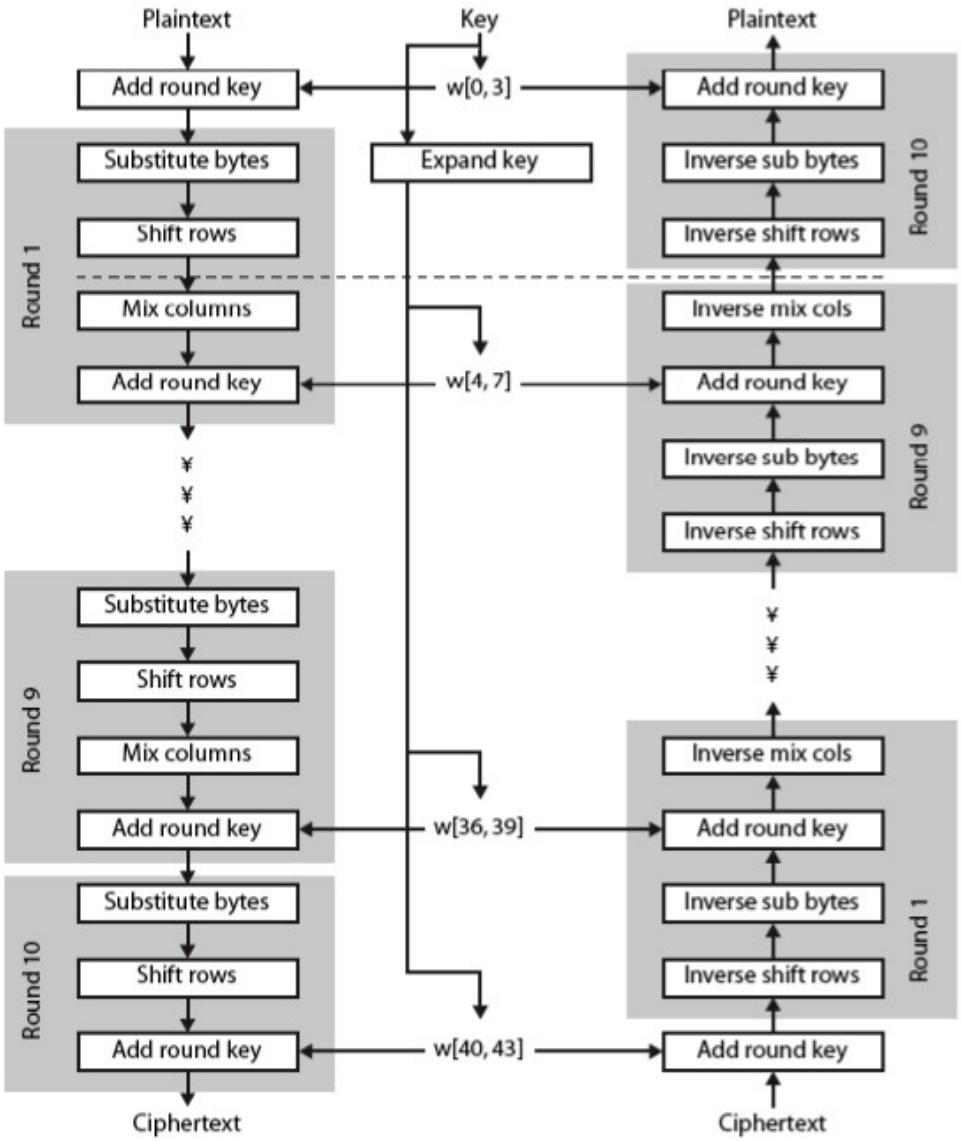
- private key symmetric block cipher
- 128-bit data, 128/192/256-bit keys
- stronger & faster than Triple-DES
- active life of 20-30 years (+ archival use)
- provide full specification & design details
- both C & Java implementations
- NIST have released all submissions & unclassified analyses

The input to the AES encryption and decryption algorithms is a single 128-bit block, depicted in FIPS PUB 197, as a square matrix of bytes. This block is copied into the State array, which is modified at each stage of encryption or decryption. After the final stage, State is copied to an output.

The key is expanded into 44/52/60 lots of 32-bit words (see later), with 4 used in each round.

The data computation then consists of an "add round key" step, then 9/11/13 rounds with all 4 steps, and a final $10^{th}/12^{th}/14^{th}$ step of byte subs + mix cols + add round key. This can be viewed as alternating XOR key & scramble data bytes operations. All of the steps are easily reversed, and can be efficiently implemented using XOR's & table lookups [17].

4.2.2.2 Basic Structure of AES [18]:



Fi

4.

Blowfish is a popular security algorithm that was developed by Bruce Schneier in the advent of the year 1994. Blowfish is a symmetric encryption algorithm, meaning that it uses the same secret key to both Encrypt and decrypt messages. The algorithm works on the same line as DES and it consumes block coding with blocks of a size of 64 bit. Blowfish became quite popular after its advent, just because Bruce Schneier himself is one of the most famous cryptology experts and above this the algorithm is non-patented, open source freely and freely available for use and modifications.

Blowfish is a 64-bit block cipher with a variable-length key. It defines 2 distinct boxes: S boxes, a P box and four S boxes. Taking into consideration the P box P is a one-dimensional field with 18 32-bit values. The boxes contain variable values; those can be implemented in the code or generated during each

initialization. The S boxes S1, S2, S3, and S4 each contain 256 32-bit values [19].

4.3 Tools to Perform Encoding and Decoding Processes:

4.3.1 ASP.NET:

In this thesis the user interface of proposed system that used to upload to or download a file from cloud was performed by ASP.NET. ASP.NET is a web framework for building modern web apps and services with .NET. ASP.NET create websites based on HTML5, CSS, and JavaScript that are simple, fast and can scale to millions of users [20].

4.3.2 Java Language:

In this thesis the service of encoding and decoding processes was performed by Java language. Java is a programming language and computing platform, first released by sun micro systems in 1995. There lots of applications and websites that will not work unless you have Java installed and more are created every day. Java is fast, secure, and reliable [21].

4.3.3 MSQL Database:

The information of encoding process was saved in MYSQL database. MSQL is the world’s most popular open source database. With its proven performance, reliability, and ease-of-use, MySQL has become the leading database choice for web-based applications, used by high profile web properties including Facebook, Twitter, YouTube, and all five of the top five websites. Additionally, it is an extremely popular choice as embedded database [22].

4.4 Method of Tools Integration:

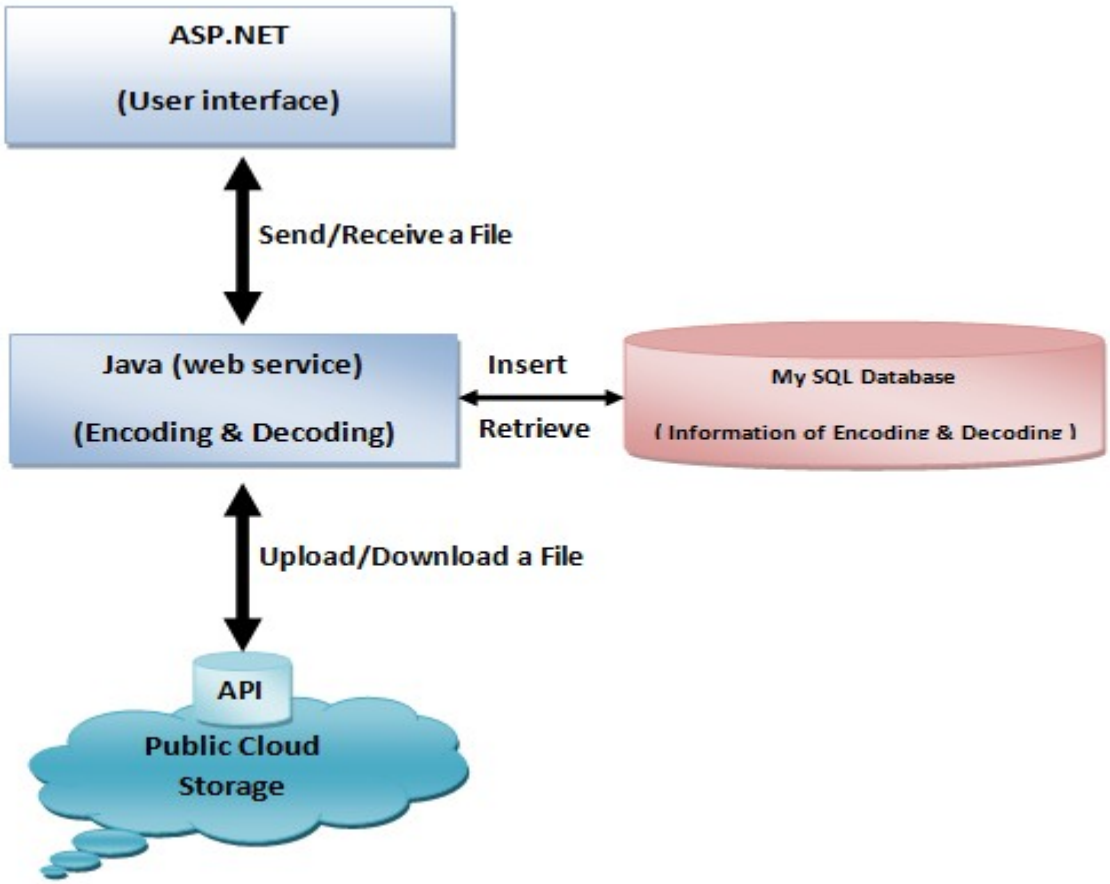


Figure 4.2: tools integration

4.5 Tool of Randomness Result:

CrypTool:

This thesis used CrypTool 1.4.41 to perform the run test. CrypTool is an open-source project. A freeware Program with graphical user interface a tool for applying and analyzing cryptographic algorithms with extensive online .help, it is understandable without deep crypto knowledge [23]

CrypTool implements more than 400 algorithms. Users can adjust these with own parameters. The graphical interface, online documentation, analytic tools and algorithms of CrypTool introduce users to the field of cryptography. CrypTool contains most classical ciphers, as well as modern symmetric and asymmetric cryptography including RSA, ECC, digital signatures, hybrid encryption, homomorphic encryption, and Diffie–Hellman key exchange.

Chapter Five

Results and Discussion

5.1 Introduction:

This chapter shows the result of file confidentiality that stored in public cloud storage after applying the proposed method on it, checking the randomness of sequence binary bits of file, and evaluate the time of encryption algorithms: AES-128, AES-256, Blowfish and proposed method.

5.2 Randomness:

All elements of the sequence are generated independently of each other, and the value of the next element in the sequence cannot be predicted, regardless of how many elements have already been produced [23].

5.3 Random Number Generation Test:

The NIST Test Suite is a statistical package consisting of 15 tests that were developed to test the randomness of (arbitrarily long) binary sequences produced by either hardware or software based cryptographic random or pseudorandom number generators. These tests focus on a variety of different types of non-randomness that could exist in a sequence [23]. The 15 tests are:

The Frequency (Monobit) Test,
Frequency Test within a Block,
The Runs Test,
Tests for the Longest-Run-of-Ones in a Block,
The Binary Matrix Rank Test,
The Discrete Fourier Transform (Spectral) Test,
The Non-overlapping Template Matching Test,
The Overlapping Template Matching Test,
Maurer's "Universal Statistical" Test,
The Linear Complexity Test,
The Serial Test,
The Approximate Entropy Test,
The Cumulative Sums (Cusums) Test,
The Random Excursions Test, and
The Random Excursions Variant Test.

5.4 The Run Test:

This thesis used runs test to check the randomness of binary sequences of file produced by proposed method [23].

Test Purpose:

The focus of this test is the total number of runs in the sequence, where a run is an uninterrupted sequence of identical bits. A run of length k consists of exactly k identical bits and is bounded before and after with a bit of the opposite value. The purpose of the runs test is to determine whether the number of runs of ones and zeros of various lengths is as expected for a random sequence. In particular, this test determines whether the oscillation between such zeros and ones is too fast or too slow.

5.5 Proposed System Implementation:

5.5.1 Front-end Implementation:

The user interface which is used for implementation is Visual Studio .NET 2013, which is implemented using ASP.NET. The user interface enabling user to upload and download file of subject asshown in figure 5-1, uploading and downloading processes are done byusing API v3that obtained from Google Drive.

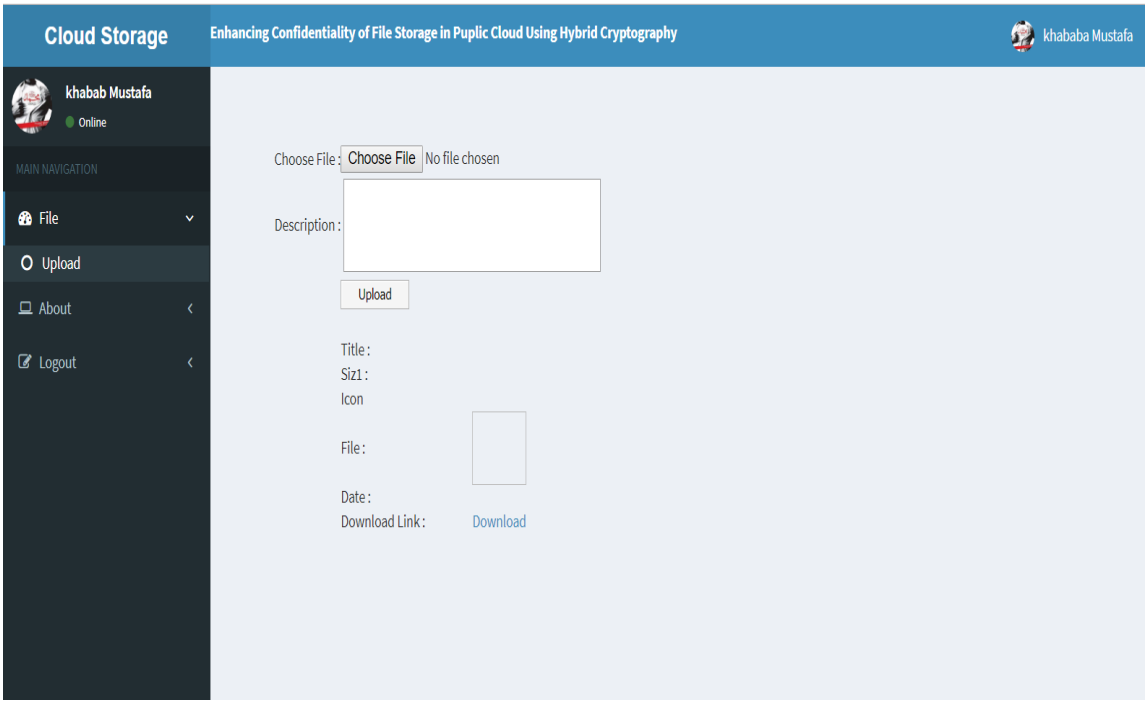


Figure 5-1: User Interface

5.5.2 Back-end Implementation:

The back-end implemented using Java programing language under NetBeans IDE-8and used MSQL as a local database. The back-end receives the file in order to encrypt uploaded file and decrypt downloaded file. The back-end included service perform the following:

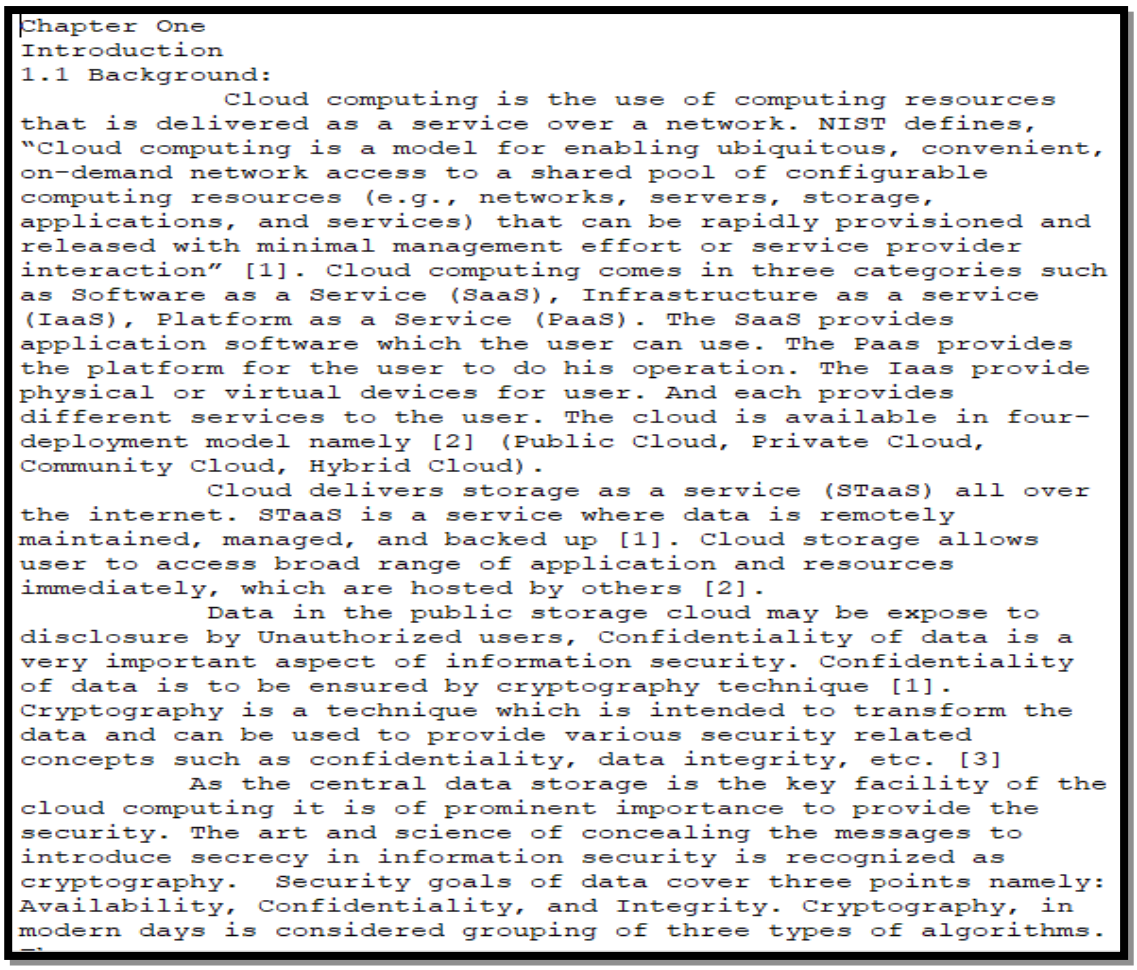
- Splitting File
- Encoding Process

- Saving the Information of Encoding Process
- Merging Blocks
- Decoding process

5.5.2.1 Splitting File:

The back-end services a file and splitting it into three blocks with different sizes randomly, the following snapshots show the original files as shown in figure 5-2, and blocks of file after splitting process, as shown in figure 5-3.

Original File:



```
Chapter One
Introduction
1.1 Background:
            Cloud computing is the use of computing resources
that is delivered as a service over a network. NIST defines,
"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" [1]. Cloud computing comes in three categories such
as Software as a Service (SaaS), Infrastructure as a service
(IaaS), Platform as a Service (PaaS). The SaaS provides
application software which the user can use. The Paas provides
the platform for the user to do his operation. The Iaas provide
physical or virtual devices for user. And each provides
different services to the user. The cloud is available in four-
deployment model namely [2] (Public Cloud, Private Cloud,
Community Cloud, Hybrid Cloud).
            Cloud delivers storage as a service (STaaS) all over
the internet. STaaS is a service where data is remotely
maintained, managed, and backed up [1]. Cloud storage allows
user to access broad range of application and resources
immediately, which are hosted by others [2].
            Data in the public storage cloud may be expose to
disclosure by Unauthorized users, Confidentiality of data is a
very important aspect of information security. Confidentiality
of data is to be ensured by cryptography technique [1].
Cryptography is a technique which is intended to transform the
data and can be used to provide various security related
concepts such as confidentiality, data integrity, etc. [3]
           As the central data storage is the key facility of the
cloud computing it is of prominent importance to provide the
security. The art and science of concealing the messages to
introduce secrecy in information security is recognized as
cryptography.  Security goals of data cover three points namely:
Availability, Confidentiality, and Integrity. Cryptography, in
modern days is considered grouping of three types of algorithms.
```

Figure 5-2: Original File before Splitting Process

Block (1) Block (2) Block (3)

Figure 5-3: Blocks of Original file After Splitting Process

5.5.2.2 Encryption Process:

The back-end selects algorithms randomly to encrypt the, the following snapshots show the blocks after encryption process, as shown in figure 5-4.

Cipher Block (1) Cipher Block (2) Cipher Block (3)

Figure 5-4: Blocks after Encoding Process

5.5.2.3 Saving the Information of Encoding Process:

The information's of encoding process (blocks sizes, file id, algorithm of each block and path of each block) saved in MSQL database as shown in figure 5-5, and this information's will be used to decrypt downloaded file.

Files Table:

fileID	fileName	size	path
45	aes.txt	236	C:\Users\khabab\Desktop\aes.txt
46	aes.txt	236	C:\Users\khabab\Desktop\aes.txt

Blocks Table:

blkID	fileID	blkSize	alg	blkPath	sequence
147	45	40	AES-256	E:/cloud/Encryption/PlainBlocks/name3.txt	2
148	45	24	AES-128	E:/cloud/Encryption/PlainBlocks/name2.txt	1
149	45	192	Blowfish	E:/cloud/Encryption/PlainBlocks/name1.txt	0
150	46	144	AES-128	E:/cloud/Encryption/PlainBlocks/name2.txt	1
151	46	96	Blowfish	E:/cloud/Encryption/PlainBlocks/name3.txt	2
152	46	24	AES-256	E:/cloud/Encryption/PlainBlocks/name1.txt	0

Figure 5-5: Tables of Encoding Process Information's

The result shown in figure 5-5 confirmed that is the file is more protection against cryptanalytic attacks (Known-plaintext attack andchosen-plaintext attack), in case of Known-plaintext attack and chosen-plaintext attack, given or chosen plaintext is not enough to get a new cipher text, because as shown in figure 5-5 the file (aes.txt) was split two times by different sizes (40, 24, 192) in first time [file ID 45] and (144, 96, 24) in second time [file ID 46], where these sizes were determined randomly,not only that but also blocks were encrypted by different algorithms with random sequence, and this makes the behavior and mode of a new cipher text undefined.

5.5.2.4 Merging Cipher Blocks:

The cipher blocks will be merged into one outsource after encoding process, the following snapshots show the original file and its encrypted form, as shown in figure 5-6:

Cipher text:

Figure 5-6: File after Merging Cipher Blocks

The result in figure 5-6 depicts how file appear after merging process,it also showed the file is more protection against(Brute-Force Attack and Cipher text-only attack),which is very difficult to predict three keys of algorithms to decrypt a file,even if one of keys was obtained (for example key of Blowfish), only part of the file will be decrypted, not fully decrypted.

5.5.2.5 Decoding Process:

- Re splitting Encrypted File:

After downloading a file from cloud storage, the back-end of software gets the blocks sizes from MySQL Database to re split encrypted file. As shown in figure 5-7:

Cipher Block (1) Cipher Block (2) Cipher Block (3)

Figure 5-7: Cipher Blocks after Re split Encrypted File

- Decryption Process: & Merging:

The back-end gets algorithms and keys from MySQL Database to decrypt cipher blocks and merging it. As shown in figure 5-8:

Chapter One
Introduction
1.1 Background:

Cloud computing is the use of computing resources that is delivered as a service over a network. NIST defines, "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" [1]. Cloud computing comes in three categories such as Software as a Service (SaaS), Infrastructure as a service (IaaS), Platform as a Service (PaaS). The SaaS provides application software which the user can use. The Paas provides the platform for the user to do his operation. The Iaas provide physical or virtual devices for user. And each provides different services to the user. The cloud is available in four-deployment model namely [2] (Public Cloud, Private Cloud, Community Cloud, Hybrid Cloud).

Cloud delivers storage as a service (STaaS) all over the internet. STaaS is a service where data is remotely maintained, managed, and backed up [1]. Cloud storage allows user to access broad range of application and resources immediately, which are hosted by others [2].

Data in the public storage cloud may be expose to disclosure by Unauthorized users, Confidentiality of data is a very important aspect of information security. Confidentiality of data is to be ensured by cryptography technique [1]. Cryptography is a technique which is intended to transform the data and can be used to provide various security related concepts such as confidentiality, data integrity, etc. [3]

As the central data storage is the key facility of the cloud computing it is of prominent importance to provide the security. The art and science of concealing the messages to introduce secrecy in information security is recognized as cryptography. Security goals of data cover three points namely: Availability, Confidentiality, and Integrity. Cryptography, in modern days is considered grouping of three types of algorithms.

Figure 5-8: Decrypted File after Decryption & Merging Process

5.6 Time Evaluation of Encoding and Decoding Processes for Encryption Algorithms: AES-256, AES128, Blowfish and Proposed Method:

This thesis analysed the performance of proposed method by measuring the computation time of the encoding and decoding process for a data file, comparing with (AES-128, AES-256, and Blowfish). In this section used file with size 1 MB, and encrypt it, the first-time using AES-128, the second time using AES-256, the third time using Blowfish, and the fourth time using proposed method.

5.6.1 Encoding Time:

The time taken to convert plaintext to cipher text is encryption time. Encryption time depends upon key size, plaintext block size and mode, encryption time must be less making the system fast and responsive. Encryption time impacts performance of the system [11].In this thesis encryption time was measured in seconds, as shown in table 5-1.

5.6.1.1 Encoding Time Results:

Encryption Algorithm	AES-128	AES-256	Blowfish	Proposed Method
Time (milliseconds)	5Sec	1.2Sec	1Sec	11Sec

Table 5-1: Encoding Time

5.6.1.2 Encoding Time Chart:

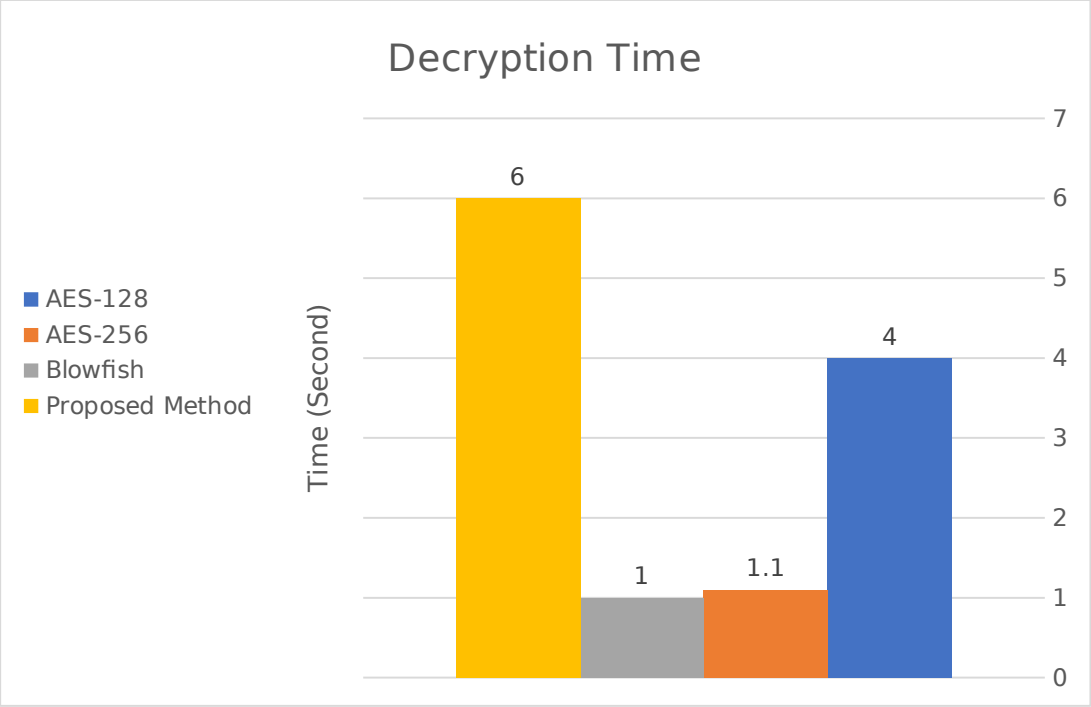


Figure 5-9: Chart of Encoding Time for AES-128, AES-256.Blowfish and Proposed Method

Figure 5-9 depict the encoding time, it was very long time comparing withAES-128, AES-256 and Blowfish, because the process of the encoding performs more operations.

5.6.2 Decoding Time:

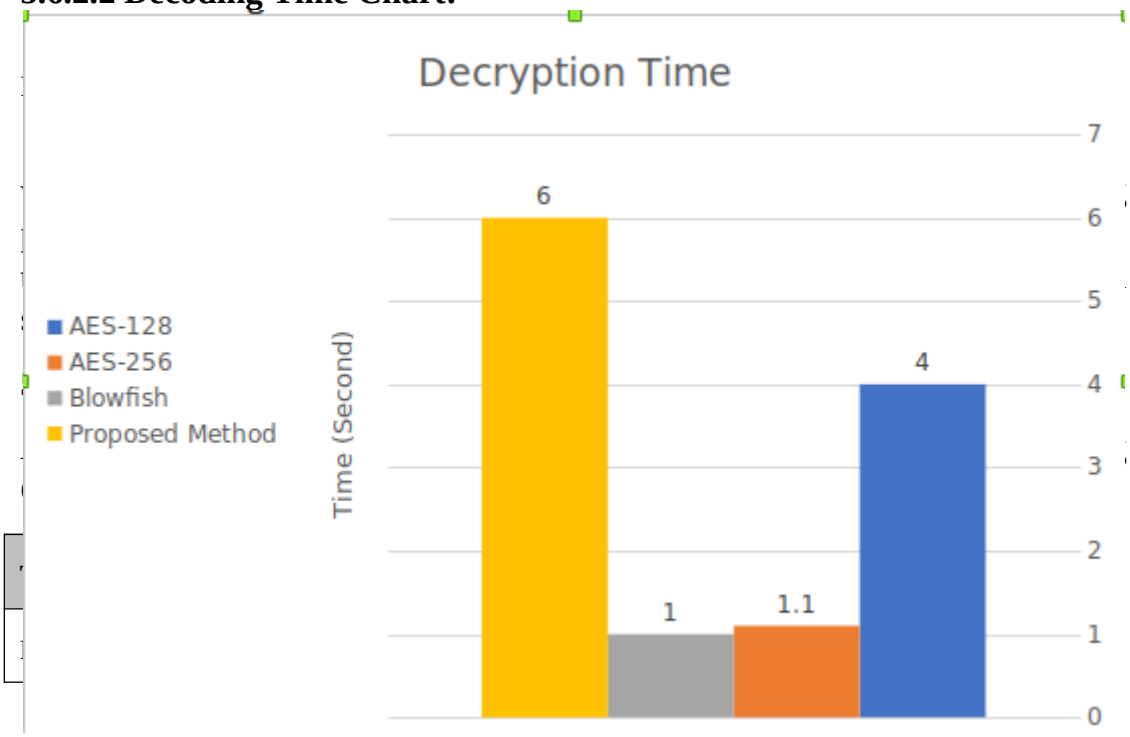
The time to recover plaintext from cipher text is called decryption time. The decryption time is desired to be less similar to encryption time to make system responsive and fast [14]. The encryption time was measured in seconds, as shown in table 5-2.

5.6.2.1 Decoding Time Results:

Decryption Algorithm	AES-128	AES-256	Blowfish	Proposed Method
Time (milliseconds)	4Sec	1.1Sec	1Sec	6 Sec

Table 5-2: Decoding Time

5.6.2.2 Decoding Time Chart:



As shown in table 5-3 the P-value obtained is ≥ 0.01 (Run Test: P-value = 10.544823), the conclusion is that the sequence is random, this result indicate to the file is more protection against Brute-Force Attack.

5.8 Discussion:

Comparison between the proposed method and AES-128, AES-256 and Blowfish algorithms based on the previous results:

- The time of encoding and decoding processes of the proposed method is longer, but achieves higher confidentiality of files.
- The cost of the proposed method is less than encryption whole file three times using mentioned algorithms.
- In the proposed method the file is more protected against cryptanalytic attacks (Known-plaintext attack andchosen-plaintext attack), because the behavior and mode of a new cipher text undefined.
- In the proposed method the file is more protected against (Brute-Force Attack and Cipher text-only attack), which is very difficult to predict three keys of algorithms.

Chapter Six

Conclusion and Recommendations

6.1 Conclusion:

Cloud Storage provides cost-effective services to individual users as well as organization. It provides huge amount of space to outsource the data to the cloud. Organization and enterprises do not possess full infrastructure to maintain their data with their premises. Data outsourcing helps to effectively maintain their data in cloud storage. Whenever user moves their data to the cloud, there are many possibilities to attack the data at rest as well as transit. Security goals of data cover three points namely: Availability, Confidentiality, and Integrity [2]. This thesis proposed technique based on providing file confidentiality using hybrid encryption algorithms. When file migrate to the cloud is fully controlled by cloud service provider not by the data owner, somewhen the provider side of cloud services exposed to attack. In this thesis to enable confidentiality of file, each user wants to upload to, or download a file from the cloud storage which must be passed to the software. The software performs encoding process before uploading and decoding process after downloading file from the cloud storage, where in this thesis the file was splitting into three blocks with different sizes and algorithms (AES (128), AES (256) and Blowfish) are chosen randomly to encrypt each block of file and saving the information's of encoding process in local database (MySQL Database) that will be needed in decoding process.

The thesis used Java programing language to implement the software that performs the encoding and decoding process, and used ASP.NET to implement the user interface for upload and download a file. The thesis showed the results of file confidentiality that stored in public cloud storage after applying the proposed method on it, the thesis discussed the randomness of sequence binary bits of file, and time of the encoding and decoding processes of this proposed method was evaluated and compared it with time of encryption algorithms: AES-128, AES-256, Blowfish, also this thesis discussed the cryptanalytic attacks and Brute-Force attack. The final results after applying proposed method had become is very difficult to predict three keys of algorithms to decrypt a file, and file had become more protection against cryptanalytic attacks and Brute-Force attack.

6.2 Recommendations:

Based on findings of the thesis, here are several recommendations to be considered:

- Reduce the time of encoding and decoding processes to enhance the performance of the system.
- In this thesis the information's of encoding process saved in the client side before upload a file to the cloud storage, which client must be downloading a file from the same computer device. Therefore, I recommend to use public server to perform encoding and decoding

processes and saving the information's of encoding process on this server instead of the client side, and then create secure channel between client and server (encoding & decoding server), which is each user wants to upload to, or download a file from the cloud storage must be passed to the server to performs encoding process before uploading and decoding after downloading from the cloud storage, as shown in figure 5-1:

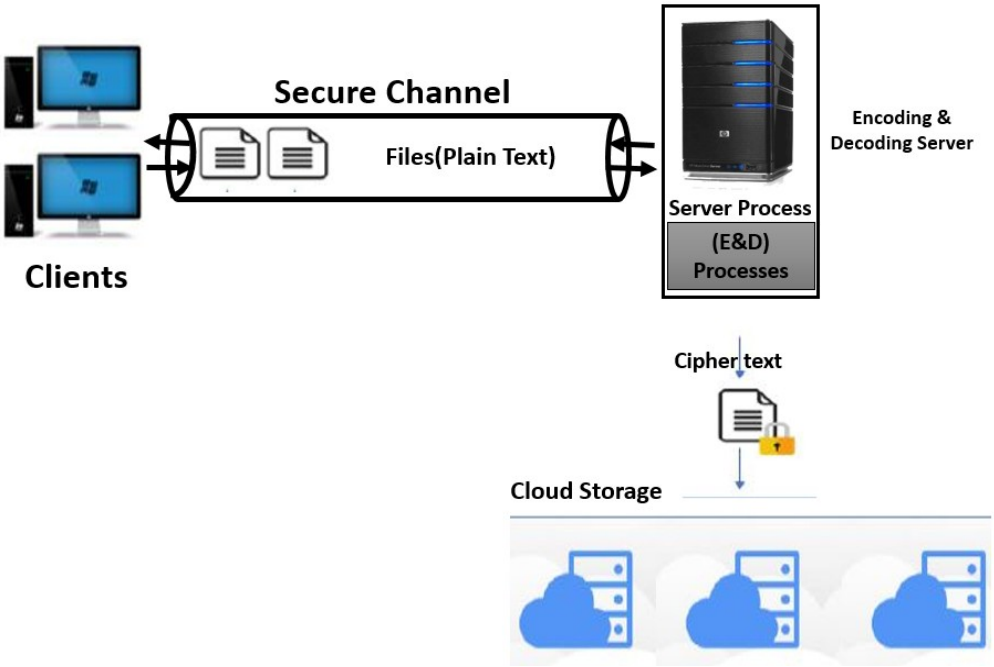


Figure 6-1: Recommend Server for Encoding and Decoding Processes

- The integrity of file should be checked after downloading from cloud storage. Because if there is any alteration in the file content, the decoding process in this thesis will not be done correctly.

References:

[1] D.I. George Amalarethinam and B. Fathima Mary, "Confidentiality Technique for Enhancing Data Security in Public Cloud Storage using Data Obfuscation",I J C T A, 2016.

[2] Jeet vyas and Prof: Prashant modi, "Providing Confidentiality and Integrity on Data Stored in Cloud Storage by Hash and Meta-data Approach ", International Journal of Advance Research in Engineering, Science & Technology, Volume 4, May-2017.

[3] Zoran Hercigonja, druga Gimnaziha and Voradzin, "Comparative Analysis of Cryptographic Algorithms ", International Journal of DIGITAL TECHNOLOGY & ECONOMY, Volume 1, 2016.

[4] Vishal R. Pancholi and Dr.Bhadresh P.Patel," Enhancement of Cloud Computing Security with Secure Data Storage using AES ", IJIRST International Journal for Innovative Research in Science & Technology, Volume 2, February 2016.

[5] Priya dhir and Sushil Garg, " Survey on Cloud Computing and Data Masking Techniques ", International Journal of Innovations & Advancement in Computer Science, Volume 6, April 2017.

[6] Jaydip Sen, " Security and Privacy Issues in Cloud Computing", Kolkata, INDIA.

[7] Monika Agrawal and Pradeep Mishra, " A Comparative Survey on Symmetric Key Encryption Techniques ", International Journal on Computer Science and Engineering (IJCSE), Vol. 4 No. 05 May 2012.

[8] url: www.searchsecurity. techtarget.com, last access in 1-9-2018.

[9] M. Saraswathi and T. Bhuvaneswari, " A Secured Storage using AES Algorithm and Role Based Access in Cloud ", IJSRSET, Volume 3, 2017.

[10] Jean Raphael Ngnie Sighom and Pin Zhang and Lin You, " Security Enhancement for DataMigration in the Cloud ", future-internet, 22 June 2017.

[11] B.Swathi and Sri .Dr. Bhaludra Raveendranadh, " Secure File Storage in Cloud Computing Using Hybrid Cryptography Algorithm ", International Journal of Advance Research in Computer Science and Engineering, Volume 6, 11 November 2017.

[12] R. Josephius Arunkumar and R. Anbuselvi, " Enhancement of Cloud Computing Security in Health Care Sector ", International Journal of Computer Science and Mobile Computing, Volume 6, August 2017.

[13] Shaikh Sufiya, Prof.Asmita Deshmukh, Prof.Ankit Sanghvi and Prof.Ashwini Sagar, " Secure Data Storage over Cloud Using Content Features Processing ", International Journal Of Engineering And Computer Science, Volume 6 October 2017.

[14] Priyadarshini Patil, Prashant Narayankar, Narayan D G and Meena S M, " A Comprehensive Evaluation of Cryptographic Algorithms: DES, 3DES, AES, RSA and Blowfish ", International Conference on Information Security & Privacy, December 2015.

[15] Jonathan Katz and Yehuda Lindell, "Introduction to modern cryptography", 2008.

[16] Ako Muhammad Abdullah, "Advanced Encryption Standard (AES) Algorithm to Encrypt and Decrypt Data", ResearchGate, June 2017.
[17] Ms NehaKhatri and Prof. V. K Kshirsagar, "Blowfish Algorithm", IOSR Journal of Computer Engineering, Volume 16, Mar-Apr 2014.
[18] https://www.asp.net, last access in 4-1-2019.
[19] https://www.java.com, last access in 4-1-2019.
[20] https://searchmobilecomputing. techtarget.com/definition/Google-Drive, last access in 5-1-2019.
[21] https://www.oracle.com/mysql/, last access in 5-1-2019
[23] https://en.cryptool.org last access in 5-1-2019.
[24] AndrewRukhin,JuanSoto, et al, " A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications", " NIST", Revised: April 2010.
[25] https://aws.amazon.com/s3/ last access in 30-1-2019.

Appendix:

- **The Run Test Example:**

Test Description:

Note: The Runs test carries out a Frequency test as a prerequisite.

(1) Compute the pre-test proportion π of ones in the input sequence: $\pi = \frac{\sum_j \varepsilon_j}{n}$.

For example, if ε = *1001101011*, then n=10 and π = *6/10 = 3/5*.

(2) Determine if the prerequisite Frequency test is passed: If it can be shown that $|\pi - 1/2| \geq \tau$, then the Runs test need not be performed (i.e., the test should not have been run because of a failure to pass test 1, the Frequency (Monobit) test). If the test is not applicable, then the *P-value* is set to 0.0000. Note that for this test, $\tau = \frac{2}{\sqrt{n}}$ has been pre-defined in the test code.

For the example in this section, since $\tau = 2/\sqrt{10} = 0.63246$, then $|\pi - 1/2| = |3/5 - 1/2| = 0.1 < \tau$, and the test is not run.

Since the observed value π is within the selected bounds, the runs test is applicable.

(3) Compute the test statistic $V_n(obs) = \sum_{k=1}^{n-1} r(k)+1$, where $r(k)=0$ if $\varepsilon_k=\varepsilon_{k+1}$, and $r(k)=1$ otherwise.

Since ε = 1 00 11 0 1 0 11, then

$$V_{10}(obs)=(1+0+1+0+1+1+1+1+0)+1=7.$$

(4) Compute $P\text{-}value = erfc\left(\frac{|V_n(obs) - 2n\pi(1-\pi)|}{2\sqrt{2n}\pi(1-\pi)}\right)$.

$$\text{For the example, } P\text{-}value = erfc\left(\frac{7 - \left(2 \cdot 10 \cdot \frac{3}{5}\left(1-\frac{3}{5}\right)\right)}{2 \cdot \sqrt{2 \cdot 10} \cdot \frac{3}{5} \cdot \left(1-\frac{3}{5}\right)}\right) = 0.147232.$$

Decision Rule (at the 1% Level)

If the computed *P-value* is < 0.01, then conclude that the sequence is non-random. Otherwise, conclude that the sequence is random.

Since the P-value obtained in step 4 of Section 2.3.4 is ≥ 0.01 (i.e., P-value = 0.147232), the conclusion is that the sequence is random.