A Research Paper on Cryptography Encryption and Compression Techniques

Asif Iqbal[1], Monalisa Nandi[2]
[1](Computer Science & Engineering , Institute of Engineering & Management, Kolkata , India, asifsubho@gmail.com)
[2](Computer Science & Engineering , Institute of Engineering & Management, Kolkata , India, monalisa80nandi@gmail.com)

Abstract— *In computing, data is information that has been translated into a form that is efficient for movement or processing. Security is the protection of assets from the theft of or damage to computer systems. Data security refers to protective digital privacy measures that are applied to prevent unauthorized access to computers, database and websites.Cryptography is an important aspect when we deal with network security.It is a very popular way to send important data in a secret way.There are many techniques of cryptography and AES is the most powerful technique. The present scenario of security system includes authenticity,integrity and confidentiality.Security is a crucial issue in now-a-days. It is about integrity,confidentiality, authentication during access or editing of internal data. Data compression is a reduction in the number of bits needed to represent data.*

Keywords— Cryptography,Data Encryption and Decryption,Security,Compression,Authentication,Confidentiality,Cipher.

1. INTRODUCTION:

Cryptography is the science of using mathematics to encrypt and decrypt data. Cryptography enables you to store sensitive data or transmit it across insecure networks like the internet so that it cannot be read by anyone except the intended recipient.Compression is used to secure the data because it use less disk space, more data can be transferred via internet.It increase the speed of data transfer from disk to memory. Main security goals are confidentiality,authentication, integrity etc.Information security is a vital issue in IT organisations. To overcome this issue most of the IT organisations are moving towards cryptography to protect their valuable information.But the organisations are also facing challenges with increasing costs of storage required to make sure that there is enough storage capacity to fulfil the present and future demand.Compressing data before encryption not only makes for shorter messages to be transmitted or stored,but also improves the security by reducing the redundancy in the plaintext and making cryptanalysis harder.Data compression is known for reducing storage and communication costs.It protect the data from eavesdropping.It tranforms a data of a given format,called plaintext,to another format,called ciphertext using an encryprion key.Now-a-days cryptography is based on mathematical theory,cryptographic algorithms are designed around computational hardness assumptions,making such algorithms hard to break in practice by any adversary.It is infeasible to break such a system by any known practical means.Cryptography is used as a weapon by organisations and government to protect the data from cyberattacks.

2. CRYPTOGRAPHY:

Cryptography is a method of protecting information and communications through the use of codes so that only those for whom the information is intended can read and process it.The pre-fix “crypt” means “hidden” or “vault” and the suffix “graphy” stands for “writing”.The root of cryptography is found in Roman and Egyptian civilization.

In computer science,cryptography refers to secure information and communication techniques derived from mathematical concepts and a set of rule based calculations called algorithms to transform messages in ways that are hard to decipher.These deterministic algorithms are used for cryptographic key generation and digital signing and verification to protect data privacy,web browsing on the internet and confidential communications such as credit card transactions and email.

In recent times,cryptography has turned into a battleground into a battleground of some of the world’s best mathematicians and computer scientists.The ability to securely store and transfer sensitive information has proved a critical factor in success in war and business.

To hide a data there mainly two techniques are used. One is Cryptography and the other is Stenography.In my paper I use Cryptography.Cryptography is the practice and study of techniques for securing communication and data in the presence of adversaries.It provides methods of converting data into unreadable form,so that Valid user can acess information at the destination.Cryptography is the technology to encrypt and decrypt the data.

3. BASIC TERMS OF CRYPTOGRAPHY:

Cryptography refers to the methodology of concealing the content of messages,the word cryptography origins from the Greek word "Kryptos",that means hidden,and "Graphikos" which means writing.

Computers are used by millions of people for many purposes such as banking,military,shopping etc.The information that we need to hide,is called plain text, It is the original text, it can be in form of characters,numerical data,pictures,program output or any kind of information.Simply, the plain text is the sending of a message in the sender side before encrypting the data, or it is the text at the receiver side after decrypting the data. The data which will be transmitted is called cipher text. It is a meaningless data so that nobody can understand except the recipient.

Cryptanalysis,on the other hand ,is the science or sometimes the art of breaking cryptosystems.Cryptography and cryptanalysis both terms are a subset of what is called as cryptology.Cryptanalysis is used to breach cryptographic security systems and gain access to the contents of encrypted messages,even if the cryptographic key is unknown.

Cipher is an algorithm which is applied to plain text to get ciphertext. It is the unreadable output of an encryption algorithm.The term "cipher" is sometimes used as an alternative term for ciphertext. Ciphertext is not understandable until it has been converted into palin text using a key.

A **Key** is an essential part of a cipher algorithm so that, in real-world ciphering, the key is kept secret, not the algorithm. Strong ciphers are designed so that, even if someone knows the algorithm, it should be virtually impossible to decipher a ciphertext without knowing the appropriate key. Consequently, before a cipher can work, both the sender and receiver must have a key or set of keys.

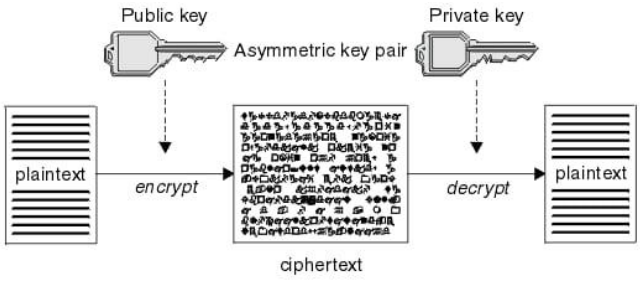

Fig:Cryptography Process using Ciphertext

Computer Security is also known as cybersecurity or IT security,is the protection of information systems from theft or damage to the hardware,the software, and to the information on them,as well as from disruption or misdirection of the services they provide. Example: Antivirus Program.

Network Security is any activity designed to protect the usability and integrity of your network and data.It includes both hardware and software technologies.Effective network security manages access to the network.It targets a variety of threats and stops them from entering or spreading on your network.The activity can be one of the following antivirus and antispyware,firewall and virtual private networks.

Internet Security is a branch of computer security specifically related to not only internet,often involving browser security and the World Wide Web,but also network security as it applies to other applications or operating systems as a whole.It is about how to prevent and detect attacks on information based systems.

4. GOALS OF CRYPTOGRAPHY:

The primary objective of using cryptography is to provide the following four fundamental information security services.The possible goals intended to be fulfilled by cryptography are:

1.Confidentiality: Confidentiality is the fundamental security service provided by cryptography. It is a security service that keeps the information from an unauthorized person. It is sometimes referred to as privacy or secrecy.Confidentiality can be achieved through numerous means starting from physical securing to the use of mathematical algorithms for data encryption.

2.Data Integrity: It is security service that deals with identifying any alteration to the data. The data may get modified by an unauthorized entity intentionally or accidently. Integrity service confirms that whether data is intact or not since it was last created, transmitted, or stored by an authorized user.Data integrity cannot prevent the alteration of data, but provides a means for detecting whether data has been manipulated in an unauthorized manner.

3.Authentication: Authentication provides the identification of the originator. It confirms to the receiver that the data received has been sent only by an identified and verified sender.Authentication service has two variants –

- **Message authentication** identifies the originator of the message without any regard router or system that has sent the message.
- **Entity authentication** is assurance that data has been received from a specific entity, say a particular website.

Apart from the originator, authentication may also provide assurance about other parameters related to data such as the date and time of creation/transmission.

4. **Non-repudiation:** It is a security service that ensures that an entity cannot refuse the ownership of a previous commitment or an action. It is an assurance that the original creator of the data cannot deny the creation or transmission of the said data to a recipient or third party.Non-repudiation is a property that is most desirable in situations where there are chances of a dispute over the exchange of data. For example, once an order is placed electronically, a purchaser

cannot deny the purchase order, if non-repudiation service was enabled in this transaction.

5. DATA ENCRYPTION AND DECRYPTION:

Encryption is the process in which a sender converts the original information to another form and sends the resulting unintelligible message out over the network. The sender requires an encryption algorithm and a key to transform the plaintext (original message) into a ciphertext (encrypted message), it's also known as enciphering. Plaintext is the data that need to be protected during transmission. The ciphertext is the scrambled text produced as an outcome of the encryption algorithm for which a specific key is used. The ciphertext is not shielded. It flows on the transmission channel. The encryption algorithm is a cryptographic algorithm that inputs plain text and an encryption key and produces a ciphertext.

In conventional encryption methods, the encryption and decryption keys are same and secret. Conventional methods are broadly divided into two classes: Character level encryption and Bit level Encryption.

- **Character-level Encryption–** In this method, encryption is performed at the character level. There are two common strategies for character-level encryption are substitutional and Transpositional.
- **Bit- level Encryption–** In this technique, firstly data (such as text, graphics, audio, video, etc.) is divided into blocks of bits, then modified by encoding or decoding, permutation, substitution, exclusive OR, rotation, and so on.

Decryption inverts the encryption process in order to convert the message back to its real form. The receiver uses a decryption algorithm and a key to transform the ciphertext back to original plaintext, it is also known as deciphering. A mathematical process utilized for decryption that generates original plaintext as an outcome of any given ciphertext and decryption key is known as Decryption algorithm. This process is the reverse process of the encryption algorithm.

The keys used for encryption and decryption could be similar and dissimilar depending on the type of cryptosystems used (i.e., Symmetric key encryption and Asymmetric key encryption).

6. SYMMETRIC KEY CRYPTOGRAPHY:

This is the simplest kind of cryptography that involves only one secret key to cipher and decipher information. Symmetrical cryptography is an old and best-known technique. It uses a secret key that can either be a number, a word or a string of random letters. It is a blended with the plain text of a message to change the content in a particular way. This key is applied to encode and decode the information. The sender uses this key before sending the message and the receiver uses it to decipher the encoded message.This is a pretty straightforward technique and as a result, it doesn't take much time. When it comes to transferring huge data, symmetrical keys are preferred.Both the sender and receiver must have a copy of the secret key.

The main disadvantage of the symmetric key cryptography is that all parties involved have to exchange the key used to encrypt the data before they can decrypt it. Modern approaches of symmetric key cryptography are executed using algorithms such as RC4, AES, DES, 3DES, QUAD, Blowfish etc.

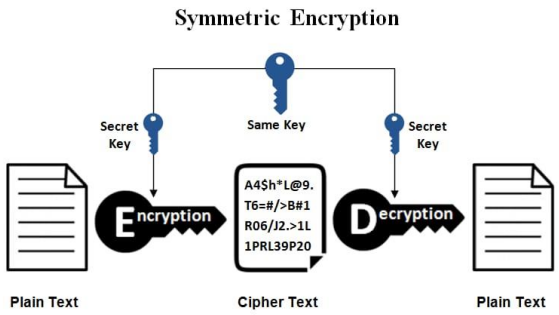

Fig: Symmetric Key Cryptography

7.ASYMMETRIC KEY CRYPTOGRAPHY:

Asymmetric Key Cryptography is a relatively new and complex mode of Data Encryption. Complex because it incorporates two cryptographic keys to implement data security. These keys are called a Public Key and a Private Key.In the two key-system is also known as public key system,one key encrypts the information and another, mathematically related key decrypts it.The Public key, as the name suggests, is available to everyone who wishes to send a message. On the other hand, the private key is kept at a secure place by the owner of the public key.By using Asymmetric Key Cryptographic method,the sender and receiver are able to authenticate one another as well as protect the secrecy of the message.

The involvement of two keys makes Asymmetric Encryption a complex technique. Thus, it proves to be massively beneficial in terms of data security. Diffie-Hellman and RSA algorithm are the most widely used algorithms for Asymmetric Key Cryptography.

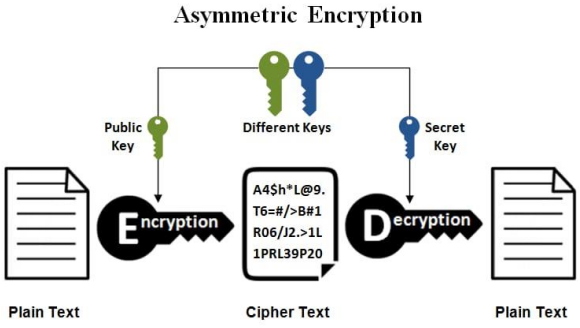

Fig: Asymmetric Key Cryptography

8. DATA COMPRESSION:

Data compression is a reduction in the number of bits needed to represent data. Compressing data can save storage capacity, speed up file transfer, and decrease costs for storage hardware and network bandwidth.
Compression is executed by a program that uses a procedure to identify how to reduce the data size.Text compression can be done by eliminating unnecessary characters, embedding a repeat character to specify repeated characters and replacing a smaller bit string for a commonly occurring bit string. Data compression can cut a text file to 50% or a considerably higher percentage of its initial size.For data transmission, compression can be done on the data content or on the whole transmission unit. When data is sent or received through the internet, larger files, may be sent in a ZIP, GZIP or other compressed format.
The purpose of compression is to make a file, message, or any chunk of data smaller. Data compression can significantly decrease the amount of storage a file takes up.The higher the compression ratio the better the compression. Because of compression, administrators save money and time on storage.Compression enhances backup storage operation and has also affected primary storage data reduction. Compression will play a significant role in data reduction as data continues to grow exponentially.

Compressing data can be a lossless or lossy process.

- **Lossless compression** enables the restoration of a file to its original state, without the loss of a single bit of data, when the file is uncompressed. Lossless compression is the typical approach with executables, as well as text and spreadsheet files, where the loss of words or numbers would change the information.
- **Lossy compression** permanently eliminates bits of data that are redundant, unimportant or imperceptible. Lossy compression is useful with graphics, audio, video and images, where the removal of some data bits has little or no discernible effect on the representation of the content.

9. DATA COMPRESSION TECHNIQUES:

- **Run Length Encoding** is the most simpliest method of compression.It is a very simple form of lossless data compression which runs on sequences having same value occurring many consecutive times and it encode the sequence to store only a single value and its count.
- **Huffman Coding** is a famous greedy algorithm that is used for lossless compression of data.It uses variable length codes are assigned to all the characters depending on how frequently they occur in the given text.The characters which occur most frequently gets the smallest code and the character which occurs least frequently gets the largest code.
- **Arithmetic Coding** is a form of entropy encoding used in lossless data compression.Arithmetic coding,which is a method of generating variable length codes, is useful when dealing with sources with small alphabets such as binary sources.It encodes data string by creating a code string which represents a fractional value on the number line between 0 and 1.Replace the entire input with a single floating-point number.
- **LZW** compression is a form of lossless compression technique.It is the compression of a file into a smaller file using a table-based lookup algorithm. It is a dictionary based algorithm that scan a file for sequences of data that occur more than once.These sequences are then stored in a dictionary and references are put where-ever repetitive data occurred.

10. CONCLUSION:

The explosive growth in the Internet, network and data security have become an inevitable concern for any organization whose internal private network is connected to the Internet. The security for the data has become highly important. User's data privacy is a central question over cloud. With more mathematical tools, cryptographic schemes are getting more versatile and often involve multiple keys for a single application. This paper presented various schemes which are used in cryptography for Network security purpose. Encrypt message with strongly secure key which is known only by sending and recipient end, is a significant aspect to acquire robust security in cloud. The secure exchange of key between sender and receiver is an important task. The key management helps to maintain confidentiality of secret information from unauthorized users. It can also check the integrity of the exchanged message to verify the authenticity. Network security covers the use of cryptographic algorithms in network protocols and network applications.
This paper briefly introduces the concept of computer security, focuses on the threats of computer network security In the future, my work can be done on key distribution and management as well as optimal cryptography algorithm for data security over clouds.

REFERENCES:

[1] William Stallings-"Cryptography and Network Security"
[2] Alfred Menzes-"Handbook of Applied Cryptography"
[3] https://en.m.wikipedia.org/wiki/Cryptography
[4] https://digitalguardian.com/blog/what-data-encryption
[5]https://www.ssl2buy.com/wiki/symmetric-vs asymmetric encryption-what-are-differences
[6] https://searchsecurity.techtarget.com/definition/cipher