International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN: 2456-3307

A Review on Compression and Encryption of Data for Secure Transmission

R. Saikumar[1], M.Sivaranjani[2]
[12]Department of Computer Science, Bharathiar University, Coimbatore, India

ABSTRACT

Data is the digital information to be stored in the computer in the form of text documents, audios, videos, or other types of data. Security is about, the protection of those assets stored in a computer. Compression algorithms can be used to reduce the redundancy of the data representation. Data compression is a very good technique which can be used to reduce the size of the data and storing the same amount of data comparatively smaller bits resulting in reducing the data storage space, resource usage. There are a lot of techniques have been implemented for the process of data compression which can be categorized as Lossy and Lossless data compression techniques. Data compression is an attractive approach to reduce the communication cost by effectively utilizing the available bandwidth in the data links. This data represents a variety of objects from the various multimedia data such as text, images, videos, sound clippings, computer programs, graphs, charts, maps, tables, etc. Over the last era, there has been recording explosion in the amount of digital data transmitted through Internet, representing text, images, video, sound, computer programs, etc. The researchers are developing the novel algorithms which can be used for data compression. It is also important to consider the security aspects of the data being transmitted while compressing it, as most of the data transmitted over the Internet is very much vulnerable to an aggregate of attacks. This presentation is focused on addressing this problem of lossless compression of multimedia files with an added security.

Keywords: Data compression, Digital Information, Security, Lossy and Lossless, Multimedia, Encryption/ Decryption

I. INTRODUCTION

Data Compression is one of the best technique for encoding the data so that it takes less storage space or less transmission time over the internet. Compression is possible because of, most of the real biosphere information's are jobless. Data Compression is a technique that can decrease the size of the data by removing unnecessary information from the file and duplicity of the file. Data compression is a skill of plummeting the number of bits needed to store or transmit the data over the network transmission. Two types of data compression techniques are there: Lossy and Lossless data compression.

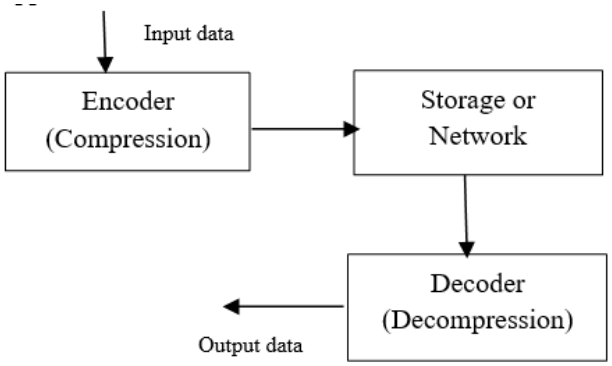

Figure 1. Diagrammatic Representation of Compression

A Lossy data compression method is one where the compressing data and then decompressing it retrieves the data that may well be different from the original. Lossy data compression is used frequently the Internet for the techniques of streaming media and telephony applications.

Lossy compression is similarly known as irreversible compression. It is a class of data encoding method, which can be used approximations and fractional data neglect to signify the content. These techniques are used to reduce the size of the data for a storage, handling the data, and transmitting the content. This data represents a variety of objects from the multimedia field such as text, images, video, sound, computer programs, graphs, charts, maps, tables, mathematical equations etc. Digital Libraries Initiative (DLI), have furnished several research projects whose goal is to collect, store, and organize an information in a digital form, and make it available for searching, retrieval, and processing via communication networks. The speed of the data transfer from the disk to memory is faster than the normal data. The security goals for the data are Confidential, Authentication, Integrity, and Non-repudiation. Information security is an emergent issue among IT organizations of all sizes. To grab this growing concern, more and more IT firms are moving towards cryptography to protect their valuable information. In addition to above concerns over securing the stored data, IT organizations are also facing challenges with ever-increasing costs of storage required to make sure that, there is enough storage capacity to meet the organization's present and future demands. It involves transforming data of a given format, called source message to data of a smaller sized format called code word. Data encryption is known for protecting an information from the spying. It transforms data of a given format, called plaintext, to another format, called ciphertext, using an encryption key. Now, the compression and encryption methods are done distinctly. Cryptography prior to the modern age was effectively synonymous with encryption, the conversion of information from a readable state to apparent nonsense. Cryptography's potential for use as a tool for espionage and sedition has led many governments to classify it as a weapon and to limit or even prohibit its use and export. Lossless compression is one of the data compression methods which can be used to allow the exceptional data to be perfectly restored from the compressed data. Lossless data compression is a method existed in many applications. For example, these methods can be used in the ZIP format and in the format of GNU tool GZIP. Some of the image files can use the formats like PNG or GIF, can be used only lossless compression, while other files like TIFF and MNG can be used either lossless or Lossy method. Maximum, lossless compression methods can ensure two things: the first step, generates a statistical model for the input data, and the second step, uses this model to map input data to bit sequences in such a way that "probable" data will produce shorter output than "improbable" data.

II. COMPRESSION

The importance of an information and the communication systems for society and the global economy is intensifying with the increasing value and quantity of data that is transmitted and stored on those systems. Here, those systems and data are also highly vulnerable to the variety of threats, such as unauthorized access, alteration of the data, and destruction of the data. Encryption and decryption is nothing but hiding the data and un-hiding the data from the unknown users due to some security reasons A cipher is used to accomplish the encryption and decryption. Merriam-Webster's Collegiate Dictionary defines cipher as —a method of transforming a text in order to conceal its meaning. An information which is veiled is called plaintext; once it has been encrypted, it is called ciphertext. To hide any data two techniques are mainly used one is Cryptography other is Steganography. In this paper we use

Cryptography. Cryptography is nothing but a method of protecting the data, which provide the countless approaches for transforming the data into an unreadable form, so that Valid User can access an information at the destination.

1) Advantages of Compression:

- Less disk space
- Reading and writing faster
- Faster file transfer
- Variable dynamic range
- Byte order independent

2) Disadvantages of Compression:

- Effects of error in transmission
- Added complication
- Slower for sophisticated method
- Unknown byte/pixel resolution
- Decompress all data

Data compression is the best method for reducing the cost of communication by encrypting the data for secure transmission using bandwidth, which is available. Compression algorithms reduce the redundancy in data representation to decrease the storage required for that data. Over the last decade, there has been an unprecedented explosion in the amount of digital data transmitted via the Internet, representing text, images, video, sound, computer programs etc. Data compression implies sending or storing a smaller number of bits. Compression is the reduction in the size of data in order to save space or transmission time. Many methods are used for this purpose, in general, these methods can be divided into two broad categories: Lossy and Lossless methods. Lossy Compression generally used for compress an image. In this original data is not identical to compressed data that means there is some loss e.g. Block Truncation Coding,Transform Coding, etc... Lossless Compression used for compress any textual data.

III. CRYPTOGRAPHY METHOD

Computers are used all over the world by people for many purposes such as banking purpose, shopping, in the military, maintaining student records, etc. Privacy is a critical issue in many of these applications, how are we need to make sure that an unauthorized party cannot read or modify messages. Cryptography is a method for the transformation of the readable and understandable data into a form which cannot be understood by the third party in order to secure the data. The information that we need to hide, is called plaintext, It's the original text, It could be in a form of characters, numerical data, executable programs, pictures, or any other information, The plaintext, for example, sending a message through our mobile can be encrypted before sending to the destination and the data can be decrypted automatically with the use of key which can be known as source code, at the time of receiving the message at the receiver side. The data which will be transmitted after encryption is called cipher text, it's refers to the string of "meaningless" data, or unclear text that nobody must understand, except the recipients. It is the data that will be transmitted exactly through the network, many algorithms are used to transform plaintext into ciphertext. Cipher is one of the method, which can be used to transform plaintext to ciphertext, this method is called encryption, and in other words, mechanism of converting readable and understandable data into "meaning less" data called encryption. Computer security is a generic term defined as a collection of tools designed to protect any data from hackers, theft, corruption, or natural disaster while allowing these data to be available to the users at the same time. The example of these tools is the antivirus program. Network security refers to any activity can be designed to protect the usability, integrity, reliability, and safety of the data during their transmission over a network, Network security deals with the hardware and software. The activity can be done by one of the following methods as anti-virus and anti-spyware, firewall, Intrusion prevention systems, and Virtual

Private Networks. Internet Security is a protocol used to protect the data during their transmission over a collection of interconnected networks, while information security is about how to prevent attacks, and to detect attacks on information-based systems. Cryptography Goals Confidentiality, Authentication, Data Integrity, Non-Repudiation, Access Control. There are two types of cryptography methods as symmetric cryptosystem and asymmetric cryptosystem which can be figured below:

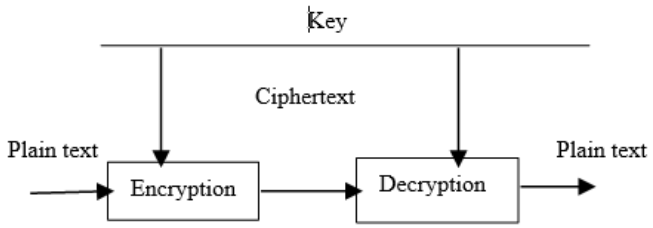

Figure 2. Symmetric Cryptosystem

In symmetric key cryptography is also known as private-key cryptography, a secret key may be held by one person or exchanged between the sender and the receiver of a message. If private key cryptography is used to send secret messages between two parties, both the sender and receiver must have a copy of the secret key.

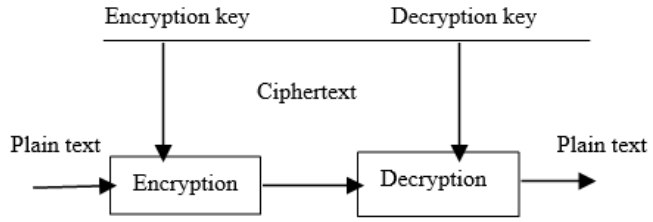

Figure 3. Asymmetric Cryptosystem

In the two-key system is also known as the public key system, one key encrypts the information and another, mathematically related key decrypts it. The computer sending an encrypted message uses a chosen private key that is never shared and so is known only to the sender. If a sending computer first encrypts the message with the intended receiver's public key and again with the sender's secret, private key, then the receiving computer may decrypt the message, first using its secret key and then the sender's public key. Using this public-key cryptographic method, the sender and receiver are able to authenticate one another as well as protect the secrecy of the message.

IV. DATA ENCRYPTION/ DECRYPTION

A. Encryption

A data encryption is a random string of bits created explicitly for scrambling and unscrambling data. Data encryption is designed with algorithms intended to ensure that every key is unpredictable and unique. Cryptography uses two types of keys: symmetric and asymmetric. Symmetric keys have the longest key, they can utilize a single key for both the encryption and decryption method of the cipher text. This type of key is called a secret key. Secret-key ciphers generally fall into one of two categories:stream ciphers or block ciphers. Data is encrypted with algorithm called encryption algorithm and encryption key. The process of encryption result in cypertext and it can be decrypted by the correct key at the receiver side. Most cryptographic processes use symmetric encryption to encrypt data transmissions but use asymmetric encryption to encrypt and exchange the secret key. Symmetric encryption is also known as the private key encryption, which can be used the same private key for both encryption and decryption methods. The risk in this system is that if either party loses the key or the key is intercepted, the system is broken and messages cannot be exchanged securely.

B. Decryption

One of the main reasons for implementing an encryption-decryption system is for data privacy. As information travels over the World Wide Web, it becomes subject to access from unauthorized individuals or organizations. Decryption is the process of taking encoded or encrypted text or other data and converting it back into text that you or the computer can read and understand. The term could be used to describe the method of un-encrypting the data manually or un-encrypting the data using the proper codes or keys. Encryption is the process of translating plain text data (plaintext) into something

that appears to be random and meaningless (ciphertext). Decryption is the process of converting cipher text back to plaintext with the presence of proper key at the receiver end.

V. CONCLUSION

Cryptography is used to ensure that the contents of a message are confidentiality transmitted and would not be altered. Confidentiality means nobody can understand the received message except the one that has the decipher key, and "data cannot be changed" means the original information would not be changed or modified.

VI. REFERENCES

[1] Swarnalata, Bollavarapu, Ruchita Sharma "Data Security using Compression and Cryptography Techniques"

[2] Manoj Patil, Prof. VinaySahu "A Survey of Compression and Encryption Techniques for SMS"

[3] Bobby Jasuja, Abhishek Pandya "Crypto-Compression System: An Integrated Approach using Stream Cipher Cryptography and Entropy Encoding"

[4] M. Burrows and D. J. Wheeler. "A Block-sorting Lossless Data Compression Algorithm", SRC Research Report 124, Digital Systems Research Cente

[5] H. Kruse and A. Mukherjee. "Data Compression

[6] Using Text Encryption", Proc. Data CompressionConference, 1997, IEEE Computer Society Press,1997, p. 447.

[7] F. Awan, Nan Zhang N. Motegi, R.Iqbal, A. Mukherjee, LIPT: A reversible Lossless Text Transformation to Improve Compression Performance., Proceedings of Data Compression

[8] The conference, Snowbird, Utah, March 2001.

[9] Dr. V.K. Govindan, B.S. Shajee Mohan "An Intelligent Text Data Encryption and Compression for High Speed and Secure Data Transmission over the Internet" CSED, L.B.S.C.E., Kasaragod, Kerala.

[10] Robert Franceschini, Amar Mukherjee" Data Compression Using Encrypted Text" o-8186-7402-4196 $5.00 0 1996 IEEE Proceedings of ADL '96

[11] Amandeep Singh Sidhu, Er. MeenakshiGarg " An Advanced Text Encryption & Compression System Based on ASCII Values & Arithmetic Encoding to Improve Data Security" International Journal of Computer Science and Mobile Computing, IJCSMC, Vol. 3, Issue. 10, October 2014, pg.45 – 51

[12] V.K. Govindan and B.S. ShajeeMohan,"IDBE - An Intelligent Dictionary Based Encoding Algorithm for Text Data Compression for High-Speed Data Transmission over Internet

[13] P.G.Howard and J.C.Vitter, Fellow IEEE" Arithmetic Coding For Data Compression".

[14] S. Kaur and V.S.Verma," Design and Implementation of LZW Data Compression Algorithm", International Journal of Information Sciences and Techniques (IJIST) Vol.2, No.4, July 2012

[15] A.Mukherjee, R.Franceschini, "Data compression Using Encrypted Text"