Dr. R. SRIDEVI *et al*, International Journal of Computer Science and Mobile Applications,
Vol.6 Issue. 4, April- 2018, pg. 222-230 ISSN: 2321-8363
UGC Approved Journal
Impact Factor: 5.515

Information Security through Compression and Cryptography Techniques

1. Dr. R. SRIDEVI Assistant professor in Department of Computer Science
PSG College of Arts & Science Coimbatore
2. SOWMIYA P R MPhil Research Scholar Department of Computer Science
PSG College of Arts & Science Coimbatore

Abstract* - *Data security is a type protecting the stored digital information. It also protects data from corruption. It is used in order to escape from the unauthorized users accessing in computers, internet websites as well as in databases of data. cryptography is a method used to encrypt and decrypt the data. A technique encryption (hiding information) data will be encoded which is used to provide secure data to prevent from fraud user and hackers where decryption is used to convert the encoded data into user readable format. It is also used for storage and transmission compression is a reduction in the number of bits needed to represent data. In this paper cryptography algorithm is user for the purpose of securing the data. It flows from compressing the data first by compression and encryption technique. When these two techniques are done at the same time data can transfer with high speed and takes less time for accessing and also for storage space.

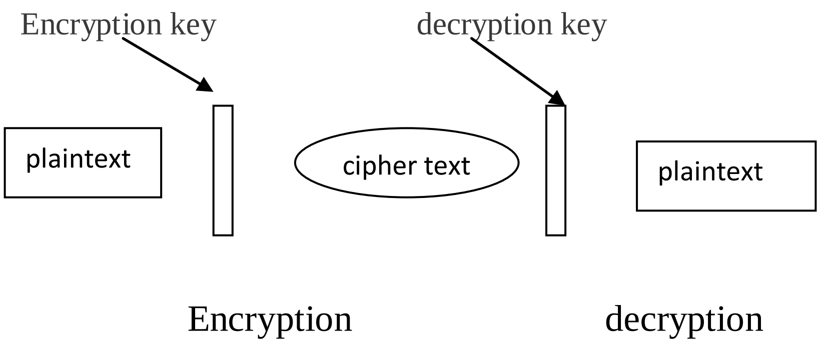
Keywords: Cryptography, Compression, Run length, Huffman, LZW, Arithmetic coding, RC4, Caesar Cipher, AES.

1. INTRODUCTION

Security is needed to guaranteeing that the data will be remain safe and no other user can access or change the information so it gives a full exactness. To secure the information, compression technique is used and the result states that it takes less space (saves money), additional information is transfer through internet. Data speed gets increased transfers from disk to memory. It prevents the compromise or loss of data contained in the database .Blocking attacks from unauthorized users or hackers. This prevents the loss of sensitive information. Data compression is the process of encoding information using fewer bits**.** It is a set of steps for packing data into a smaller space, while allowing for the original data to be seen again. Data Encryption is the process of translating plain text data (plaintext) into something that appears to be random and meaningless (cipher text). Decryption is the process of converting cipher text back to plaintext. To encrypt more than a small amount of data, symmetric encryption is used. Now in current the compression and encryption is required with more processing speed and cost. This paper combines both the compression and the encryption technique to overcome with less storage and cost.

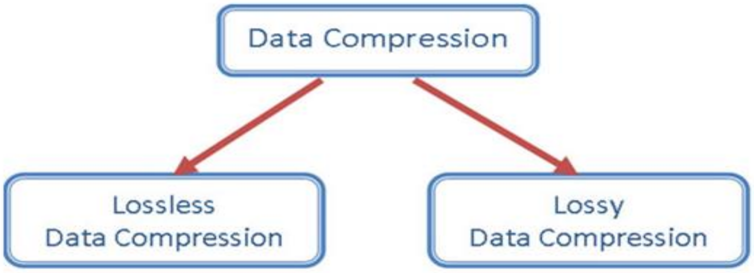
1.1 CRYPTOGRAPHY

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 art of protecting information by transforming it (encrypting it) into an unreadable format, called cipher text. Only those who possess a secret key can decipher (or decrypt) the message into plain text. Cryptography not only protects data from theft or alteration, but can also be used for user authentication.



1.2 DATA COMPRESSION

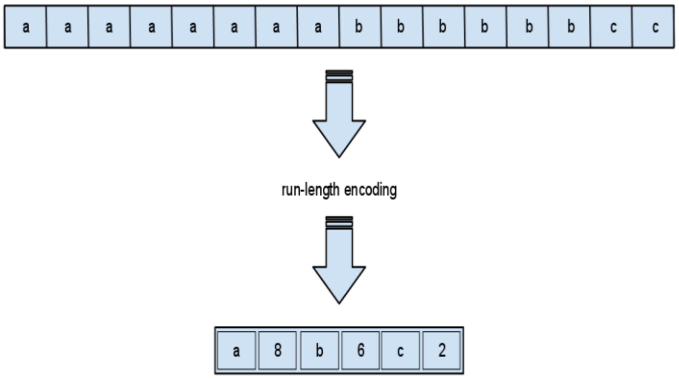
A common **data compression** technique removes and replaces repetitive data elements and symbols to reduce the data size. **Data compression** for graphical data can be lossless compression or lossy compression, where the former saves all replaces but save all repetitive data and the latter deletes all repetitive data. In this uncommon information is not vague to compacted information that recommends there is some calamity e.g. Piece Truncation Coding, Transform Coding, etc Lossless Compression utilized for pack any printed information. In this uncommon information and pounded information are measure up to that finds there is no incident e.g. Run Length Coding, Huffman Coding, LZW, Arithmetic Coding.



2. COMPRESSION METHODOLOGIES

2.1 RLE (run length encoding)

Run length encoding 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 counts. RLE works by reducing the physical size of a repeating string of characters. This repeating string, called a *run*, is typically encoded into two bytes. The first byte represents the number of characters in the run and is called the *run count*.



2.2 Huffman coding

It is used to compress or encode data. Normally, each character in a text file is stored as eight bits (digits, either 0 or 1) that map to that character using an encoding called ASCII. A Huffman-encoded file breaks down the rigid 8-bit structure so that the most commonly used characters are stored in just a few bits ('a' could be "10" or "1000" rather than the ASCII, which is "01100001"). The least common characters, then, will often take up much more than 8 bits ('z' might be "00100011010"), but because they occur so rarely, Huffman encoding, on the whole, creates a much smaller file than the original.

2.3 Arithmetic coding

Arithmetic coding is a common algorithm used in both lossless and lossy data compression algorithms. It is an entropy encoding technique, in which the frequently seen symbols are encoded with fewer bits than rarely seen symbols. It has some advantages over well-known techniques like Huffman coding.

2.4 Variable length encoding

Variable-length codes can allow sources to be compressed and decompressed with *zero* error (lossless data compression) and still be read back symbol by symbol. With the right coding strategy an independent and identically-distributed source may be compressed almost arbitrarily close to its entropy. This is in contrast to fixed length coding methods, for which data compression is only possible for large blocks of data, and any compression beyond the logarithm of the total number of possibilities comes with a finite (though perhaps arbitrarily small) probability of failure.

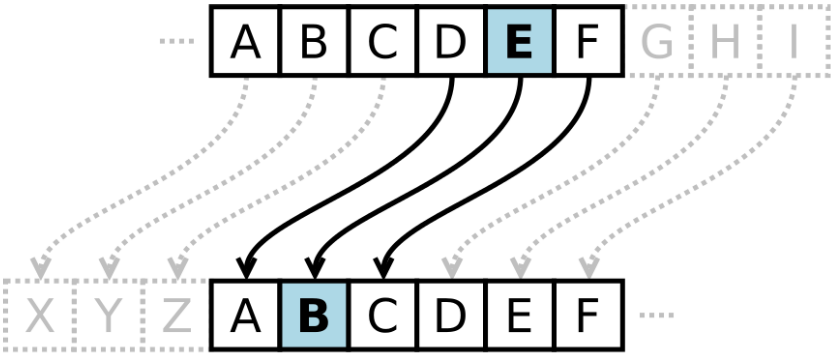
3. CRYPTOGRAPHIC TECHNIQUES

3.1 Rc4(Rivest Cipher 4)

In cryptography, **RC4** (Rivest Cipher 4 also known as **ARC4** or **ARCFOUR** meaning Alleged RC4, see below) is a stream cipher. While remarkable for its simplicity and speed in software, multiple vulnerabilities have been discovered in RC4, rendering it insecure. It is especially vulnerable when the beginning of the output keystream is not discarded, or when nonrandom or related keys are used.

3.2 Ceaser cipher

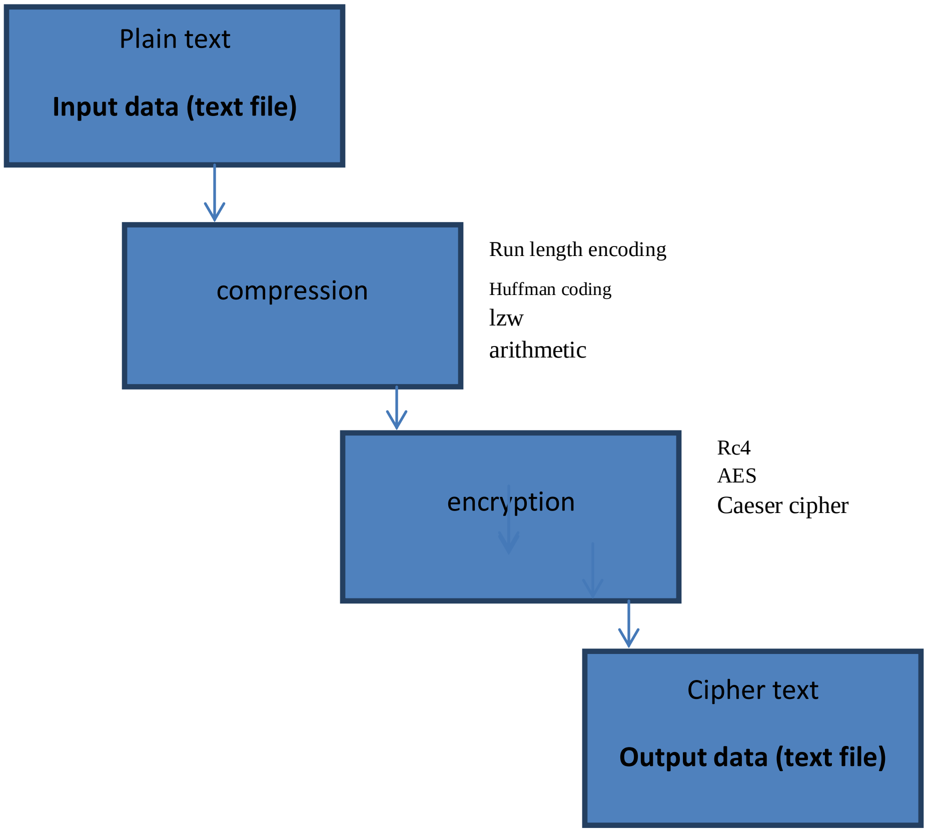
The **Caesar Cipher**, also known as a shift cipher, is one of the oldest and simplest forms of encrypting a message. It is a type of substitution cipher where each letter in the original message (which in cryptography is called the plaintext) is replaced with a letter corresponding to a certain number of letters shifted up or down in the alphabet. For example, with a shift of +3 (to right) word "B" will become "E".



3.3 AES

Advanced Encryption Standard (AES) is a cipher, meaning that it is a method or process used to change raw information (usually human readable) into something that cannot be read. This part of the process is known as encryption. The method uses a known, external piece of information, called a key, to uniquely change the data. An example might be your computer login password, or the password to your account on a bank machine. Further, the process is reversible, meaning that it can be applied again to put the information back into its original form. This part is known as decryption.

4. IMPLEMENTATION



Demonstrates Flow of work of the blend of pressure furthermore, encryption. In this, four pressure and three cryptographic systems are connected on content information at that point as certain execution investigation concerning document estimate, their pressure proportion and their execution time. Stream of work is to take any content record first we pack a content document after that whatever yield will understand that will be go to encryption part to encode that record.

Mix of pressure and encryption calculations are connected on content record are as per the following:
Four pressure methods which are RLE (Run Length Encoding), Huffman coding, LZW, Arithmetic and three cryptographic methods which are RC4, Caesar Cipher and AES are utilized. To secure the information more we utilize mix of pressure and cryptographic methods on content document.
So here the mix which are utilized to secure information are: RLE + (RC4 and Caesar Cipher), Huffman + (RC4 and AES), LZW + (RC4 and AES), Arithmetic + (RC4 and AES) with their comparing document measure, pressure proportion and execution time.

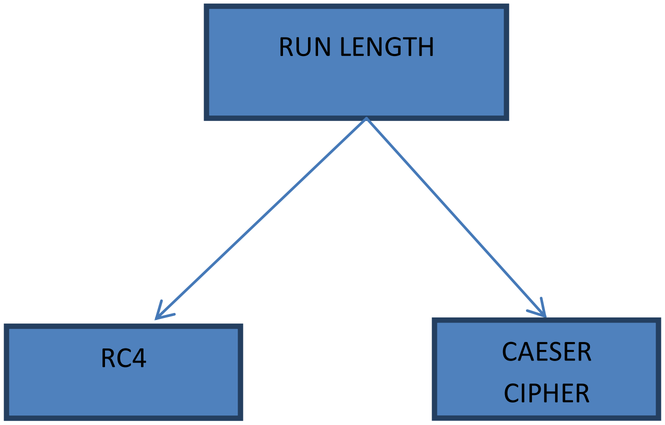


Fig : RLE + (RC4/Caesar Cipher)

Run length encoding with RC4 and Caesar Cipher will apply on five distinctive size of content document and result will be done. This mix is great when two or then again more continuous character will happen in any content document. On the off chance that there are less number of back to back character in any content record at that point result will be bad as near to different methods.

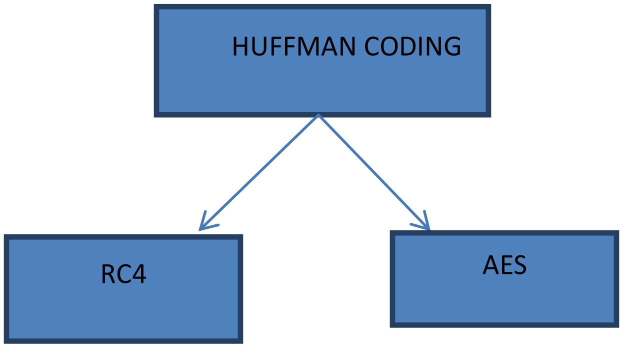


Fig : Huffman + (RC4/AES)

Huffman with RC4 and AES will apply on those five diverse size of content document. This mix is best finished any pressure methods. Their pressure proportion is exceptionally great. In the event that Huffman coding is utilized to pack a record then the document measure is simply half and it regards send the more information on web.

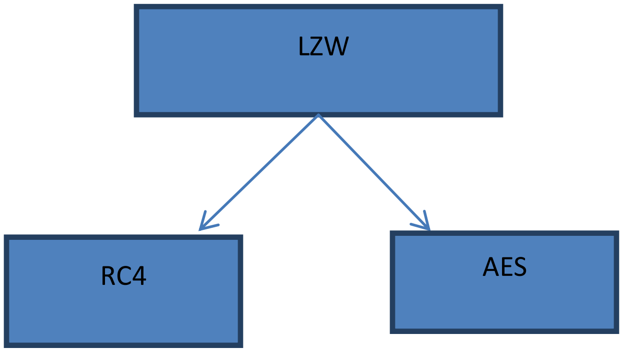


Fig : LZW + (RC4/AES)

LZW and Arithmetic with RC4 and AES will apply on those five diverse size of content document. It pack a record furthermore, give result that is great however not as a decent as Huffman will give. There are minor distinction in values after pressure of both the strategies (Arithmetic and LZW). After pressure, encryption will perform and there isn't to such an extent contrast in values.

In the wake of applying the mix following outcome will be completed. Here we take five distinctive content record which are diverse in measure. After the outcome we examination that higher the record measure better the pressure proportion. After the pressure we perform encryption to give better security. Different blend which are utilized on content record that shows which mix we will use for better security and better pressure proportion.

5. CONCLUSION

In this paper it tells about the security of the data by using encryption which translates the data into a secret code. It is the most effective way to achieve data security which access to a secret key or password that enables to decrypt it. Unencrypted data is called plain text; encrypted data is referred to as cipher text .Compressing data before encryption not only makes for shorter messages to be transmitted or stored, but also improves security. Compression is done before encryption. Cryptography can be used to provide message confidentiality and integrity and sender verification. The basic functions of cryptography are encryption, decryption and cryptographic hashing. For data security, combination of compression and cryptographic techniques are used. If encryption are done then it prevents from fraud users at the same time when compression are done then it takes less processing time and more speed.

REFERENCES:

1. “A Study on Elliptic Curve Cryptography” in International Journal on Research Innovations in Engineering Science and Technology (IJRIEST), (ISSN 2455-8540), Volume 2, Issue 7, July-2017, pp: 546 – 550.

2. T.SubhamastanRao, M.Soujanya, T.Hemalatha, T.Revathi, “Simultaneous data compression and encryption” (IJCSIT) International Journal of Computer Science and Information Technologies, ISSN 0975-9646, Volume2(5), 2011.

3. Senthil Shanmugasundaram, Robert Lourdusamy “A Comparative Study Of Text Compression Algorithms” International Journal of Wisdom Based Computing, Vol. 1 (3), December 2011.

4. “A Survey on Cryptographic Cloud Storage Techniques” in International Journal of Engineering Sciences &amp; Research Technology(IJESRT), ISSN: 2277-9655, July 2017, pp: 602-605.(Impact Factor:4.116)

5. “An Implementation of Hybrid Cryptographic Protocol for Facial Image Security” in International Journal of current research in science and technology (ISSN: 2394-5745), Volume 1 Number 8 (2015).

6. T.D.B Weerasinghe “Analysis of a Modified RC4 Algorithm” International Journal of Computer Applications, ISSN0975 – 8887, Volume 51– No.22, August 2012.

7. “Secured Transmission of Data in cloud Environment Using Elliptic curve Cryptography” in International Journal of Innovative Research in Computer and Communication Engineering (IJRCCE) (ISSN (online): 2320-9801), Vol. 3, Issue 8, August 2015(impact factor: 5.618).

8. “A Secured Asymmetric Cryptosystem for Multi-Biometric Security” in International Journal of Trend in Research and Development (IJTRD) (ISSN (online): 2394-9333), Vol. 2, Issue 5, September-October 2015(impact factor : 0.924).

9. “Protection of Private Data over the Transmission through Electronic Gadgets using RSA Algorithm” in International Journal of Computer Science and Information Security (IJCSIS)(ISSN (online): 1947-5500), Vol. 13, Special Issue , July 2015, pp.40 – 43, (Impact factor:0.74).

10. “A Comparative Study on the Performance and the Security of RSA and ECC Algorithm” in Special Issue Published in International Journal of Advanced Networking and Applications (IJANA), March 2015 Pages: 168 – 171, (impact factor: 3.462)

11. “A Novel Scheme for Remote Data Storage – Dual Encryption” in an International Journal of Research in Information Technology (IJRIT) ISSN: 2001-5569 Volume 2 Issue 4 April 2014.(Impact factor:1.738)

12. “An Approach for Secure Data Storage in Cloud Environment” in an International Journal of Computer and Communication Engineering (ISSN: 2010-3743) as one volume, and indexed by World Cat, Google Scholar, and Engineering &amp; Technology Digital Library, Ulrich&#39;s Periodicals Directory.

13. Ms. Ayushi Aggarwal, Anju “Enciphering Data for Larger Files” International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X, Volume 3, Issue 5, May 2013.

14. Ms. Ayushi Aggarwal, Anju “Enciphering Data for Larger Files” International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X, Volume 3, Issue 5, May 2013.

15. William Stallings. ----“Cryptography and Network Security Principles and Practice”

16. Atul Kahate. ------Computer and Network Security