International Journal of Innovative Research in Information Security (IJIRIS) ISSN: 2349-7017
Issue 08, Volume 4 (August 2017) www.ijiris.com

SURVEY OF COMPRESSION AND CRYPTOGRAPHY TECHNIQUES OF DATA SECURITY IN E-COMMERCE

Samar lofty
Dean, Faculty of Computer and Informatics,
Zagazig University, Zagazig, Egypt
nasserhr@gmail.com
Abdel Nasser H. Zaied
Information system department,
Faculty of Computer and Informatics,
Zagazig University, Zagazig, Egypt
Samarlotfy27@yahoo.com

Manuscript History
Number: **IJIRIS/RS/Vol.04/Issue08/AUIS10080**
DOI: 10.26562/IJIRIS.2017.AUIS10080
Received: 07, August 2017
Final Correction: 12, August 2017
Final Accepted: 17, August 2017
Published: August 2017
Citation: **lofty Samar & Nasser, H. Z. A. (2017). SURVEY OF COMPRESSION AND CRYPTOGRAPHY TECHNIQUES OF DATA SECURITY IN E-COMMERCE. IJIRIS:: International Journal of Innovative Research in Information Security, Volume IV, 01-08. doi: DOI: 10.26562/IJIRIS.2017.AUIS10080**
Editor: Dr.A.Arul L.S, Chief Editor, IJIRIS, AM Publications, India

Abstract: E-commerce is one of most important fields in business process .many organization depend on e-commerce to enhance their work .there are many challenges faces the e commerce and has a negative effects in applying process such as security risks. Any organization depend on e-commerce must take data security in their consideration. Data security contains many areas, and the most important area is to secure the data from un-authorized access .Cryptography contain many techniques which can be used to apply the security for data. Cryptography may be used for many issues such as prevent unauthorized access, data integrity and authentication. Compression techniques contain some techniques which be used to reduce the bits which represent the data. In this survey we presents some of compression and cryptography techniques which be used to improve data security.
Keywords: E-commerce, E-commerce security, data security, encryption Compression

I. INTRODUCTION

Nowadays, Electronic commerce is one of important fields in the process of development any business. Electronic commerce means buying or selling anything from the internet. The commerce activities through internet has been speeded and grown in the century. When it comes to payment, one needs to establish a sense of security. Customers must be able to select a mode of payment and the software must verify their ability to pay. This can involve credit cards, electronic cash, encryption, and/or purchase orders. The more of these techniques are supported by an E-commerce package, the more secure the system can be, and therefore the more customers are benefits from E-commerce abilities [1-4].E-commerce is changed the way of the business think such as the business operation and, practices and the relationship between the business and customers. Security issues are an important topic in e-commerce.

Any organization depends on e commerce must take in consideration about how to secure their data and customers data. E-commerce contains many shapes of data such as, databases, transaction records, commercial transactions, user data and etc. The data are very important to the parties involved in e-commerce, so we must assure their security completely [5].Nowadays ,e-commerce contains many security issues such as , Web access control, user authentication, authorization control, safety audit, backup and recovery, data encryption and etc. These issues have a great impact in the data security process. Cryptography is one of the most effective technologies of data security. In this survey we presents some of cryptographic techniques which be used in data security. Compression techniques may be used mixed with compression technique to improve the security process, so in these survey we presents some of these techniques[6-8]. The survey contains five main sections: Introduction, cryptographic technique section, which introduces different types of Cryptographic Techniques, compression techniques, hybrid techniques, conclusion and references.

II. CRYPTOGRAPHIC TECHNIQUES

Cryptography is the science of protecting data, which provides techniques of converting data into secured form to prevent unauthorized actions. Cryptography may be used to improve data integrity, authentication and enhancing access control. cryptography contain two operation ,encryption process which is related to sender ,and decryption process which is related to receiver .Figure 1 shows cryptographic process between sender and receiver .

Figure 1: Cryptography Process

2.1-Types of Cryptographic Techniques

In this subsection we presents some of cryptographic techniques which may be used in process of improving data security.

i. Data Encryption Standard

Data encryption standard is one of most widely used cryptographic techniques which is developed by Horst Fiestel and approved by NBS (National Bureau of Standards, now called NIST -National Institute of Standards and Technology) in 1978. The DES was standardized by the ANSI (American National Standard Institute) under the name of ANSI X3.92, better known as DEA (Data Encryption Algorithm)[9,10]. The Data Encryption standard is one of most important symmetric techniques which depend on the same key for encryption and decryption . DES uses 56 bits key for encryption and decryption. It completes the 16 rounds of encryption on each 64 bits block of data. Data encryption standard works on a particular principle. Data encryption standard is a symmetric encryption system that uses 64-bit blocks, 8 bits (one octet) of which are used for parity checks (to verify the key's integrity). Each of the key's parity bits (1 every 8 bits) is used to check one of the key's octets by odd parity, that is, each of the parity bits is adjusted to have an odd number of '1's in the octet it belongs to. The key therefore has a real useful length of 56 3bits, which means that only 56 bits are actually used in the algorithm. So it would take a maximum of 256 or 72,057,594,037,927,936, attempts to find the correct key.DES can be used in many application such as image processing ,network ,medical application and etc .

ii. Triple Data Encryption Standard

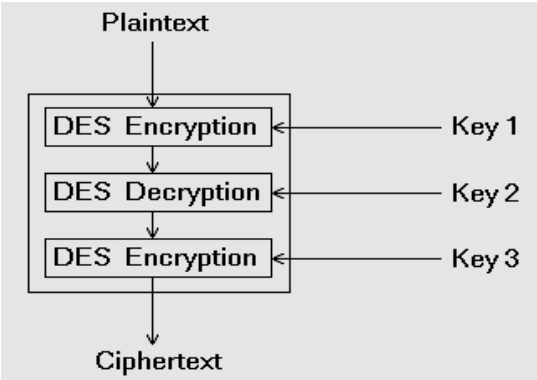

Figure 2 - 3DES Structure

Triple data encryption or 3DES is one of most important cryptography algorithm and one of most widely used algorithms. The algorithm is the same idea of traditional DES but it apply the algorithm three times to improve the complexity of the algorithm as figure 2 .triple des one of most complex algorithm and it is widely used in many fields . 3Des is preferred than the traditional DES because it is add more security and complexity. The main problem of the algorithm is the time consuming.[12,15]

iii. Rivest-Shamir-Adleman Algorithm

RSA is a public key system designed by Ron Rivest, Adi Shamir, and Leonard Adleman in 1978.The RSA algorithm contain of three stage the first stage is to generate the keys, the second stage is the encryption process and the last stage is decryption process. The RSA algorithm depends on two distinct keys p&Q which are generated using Euler theory, Chinese remainder theorem, hamming weight and exponential functions key has been generated and then encryption process takes place. Decryption has been done in the receiver section by using the public key concept. [11].the next figure show the sequence steps of RSA.

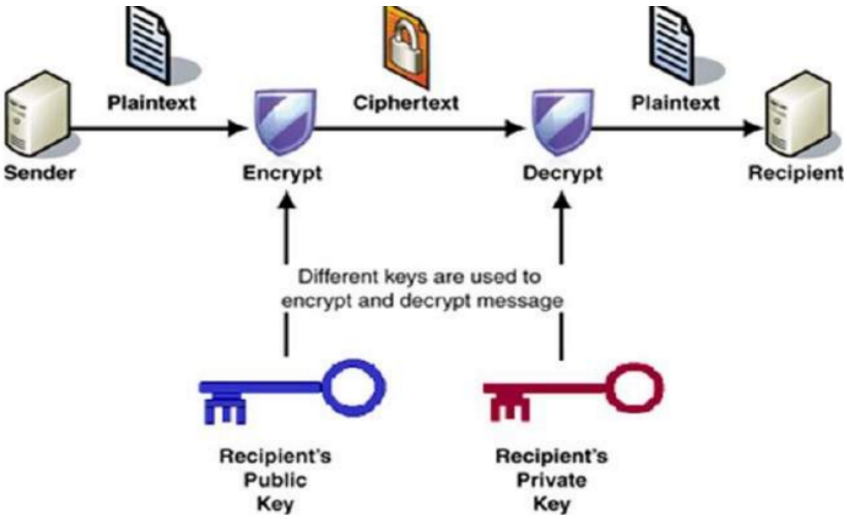

Figure 3: RSA Algorithm (Asymmetric Key Cryptography)

RSA called a public key cryptography because it depends on same key for encryption and decryption. The key is also shared between the sender and receiver. The next algorithms show the pseudo code of RSA algorithm.

Algorithm 1: pesudocode of RSA

```
Choose p and q
Compute n = p * q
Compute φ(n) = (p - 1) * (q - 1)
Choose e such that 1 < e < φ(n) and e and n are co-prime.
Compute a value for d such that (d * e) % φ (n) = 1.
compute the public key, Public key is (e, n)
Compute the private key ,Private key is (d, n)
For encryption C=me(mod n) and decryption m =cd(mod n)
```

iv. Advanced Encryption Standard

Advance Encryption standard was developed in 1997, the National Institute of Standards and Technology (NIST) announced an initiative to choose a successor to DES; AES take the place of DES and 3DES in many fields during its complexity and security.AES is symmetric encryption which is depend on different key in sender and receiver . AES depends on three blocks ciphers; the first one is AES-128, the second one is AES-192 and the third one is AES-256. Each cipher encrypts and decrypts data in blocks of 128 bits using cryptographic keys of 128 bits, 192 bits and 256 bits, respectively. The algorithm depend on internal structure in rounds ,it contains 10 rounds ,these round add more complexity and security to the algorithm .it contains 10 rounds for 128-bit keys, 12 rounds for 192-bit keys, and 14 rounds for 256-bit keys[14].each round work with itself identically except the last round .each round contain four stage ;the first one is substitution, the second stage is row shifting ,the third stage is column mixing ,the last stage is to add round keys. This stage occurs as internal structure in each round, these adds more complexity and security to algorithm.

v. Blowfish

Blowfish was developed by bruce schneier in 1993. It is basically a symmetric block cipher .the algorithm depends on variable length key 32 bits to 448 bits, the algorithm take 64 data block as input .the algorithm depend on the rounds in its process ,it contain 16 rounds.

The algorithm also uses large key dependent S-Boxes. Each S-box contains 32 bits of data. The algorithm is one of the algorithms which are developed to replace the DES algorithm. The algorithm starts by dividing data into blocks with size 64 bits and work with each block individually. Blowfish is known for both its tremendous speed and overall effectiveness as many claim that it has never been defeated. . Blowfish can be found in software categories ranging from e-commerce platforms for securing payments to password management tools, where it used to protect passwords. It's definitely one of the more flexible encryption methods available.

vi. Twofish

Two fish is one of most important symmetric algorithm which is depends on feistel structure. The algorithm is developed by bruce schneier in 1998. Twofish also uses block ciphering like Blowfish. It is efficient for software that runs in smaller processor (smart cards) and embedding in hardware. It allows implementers to customize encryption speed, key setup time, and code size to balance performance. Twofish is license-free, un-patented and freely available for use. In twofish encryption it uses key sizes of 128, 192 and 256 bits. It uses the block size of 128 bits and there are 16 rounds of encryption in this encryption algorithm.

vii. RC5

The RC5 encryption algorithm was designed by Professor Ronald Rivest
The RC5 encryption algorithm was designed by Professor Ronald in December 1994 [15]. RC5 is a symmetric-key block cipher. AES (Advanced Encryption Standard) is directly based on RC5. It uses key sizes 0 to 2040 bits but suggested count is 128 bits. RC5 uses block sizes of 32, 64 or 128 bits but 64 bits are suggested. It is fiestel-like network [16]. It has 1 to 255 encryption rounds but 12 rounds are suggested originally. It is suitable for hardware and software implementation, because it uses only those operations which are available in typical microprocessor [15].the next figure shows the sequence structure of RC5 algorithm.

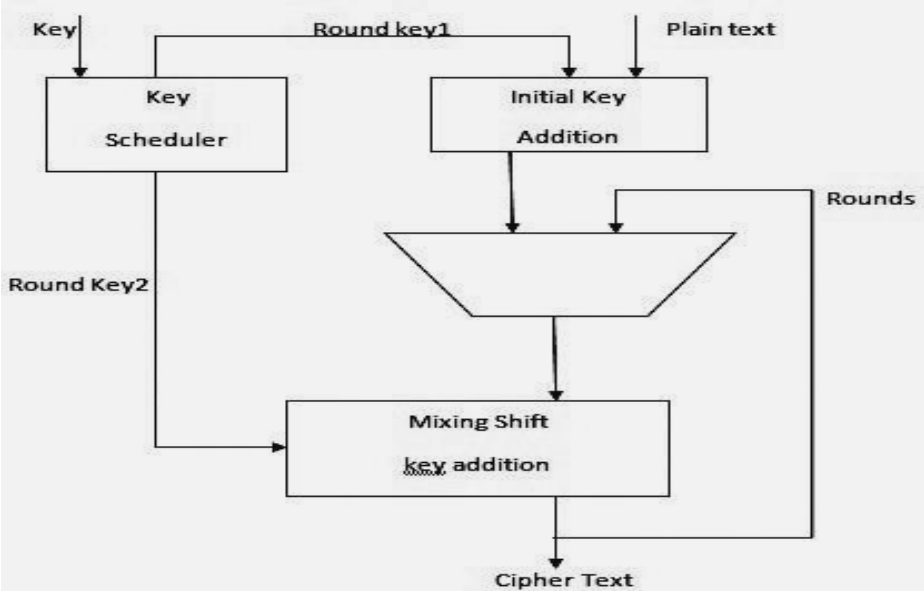

Figure 4:RC5 Encryption Procedure

RC5 has attracted the attention of many researchers in the cryptographic
The next table shows the main parameters of the previous encryption algorithms

Algorithms	Advantages	Limitation	References
DES	it's better than XOR, and probably better than some crypto scheme you thought up yourself.	There is no strong limitation found rather than its small key size which offers less security. The only successful attack on DES is Brute force attack. It's another weak point is its encryption speed which is very slow	[9],[10]
3DES	The main advantage of Triple DES is that it is three times secure (as it is combination or three DES algorithms with different keys at each level) than DES that's why it is preferred over simple DES encryption algorithm. It provide adequate security to the data	it consumes lot of time and its encryption speed also less than DES encryption algorithm.	[15],[12]

RSA	It provides good level of security	The main disadvantage is its encryption speed. It consumes lot of time to encrypt data. Actually this is disadvantage of asymmetric key algorithms	[14]
AES	AES is more secure (it is less susceptible to cryptanalysis than 3DES). AES supports larger key sizes than 3DES's 112 or 168 bits.AES is faster in both hardware and software		[11],[12],[43]
Blowfish	Blowfish is yet another algorithm designed to replace DES. Blowfish can be found in software categories ranging from e-commerce platforms for securing payments to password management tools, where it used to protect passwords. It's definitely one of the more flexible encryption methods available	Blowfish has disadvantage over other algorithms in terms of time consumption.	[13],[44]
Twofish	It allows implementers to customize encryption speed, key setup time, and code size to balance performance Twofish is as fast as DES on throughput		[9]
RC5	It is suitable for hardware and software implementation,because it uses only those operations which are available in typical microprocessor		[9],[15],[16]

III. COMPRESSION TECHNIQUES

Compression techniques one of the most important techniques which is being used in many application such as data transmission, operating system and etc .there are two types of compression the first one is lossless compression which the data after compression and decompression are the same ,and the second one is lossy compression which data loss some bits . [23].in the next two sub section we introduces some of lossless and lossy compression.

3.1 Examples of lossless Compression Techniques:

i. Run Length Encoding

Run Length Encoding (RLE) is one of simplest compression techniques .run length coding is depending on replacing the sequence number of characters by the number of sequence and single character only. [26]. Run-length encoding is also often used as a preprocessor for other compression algorithms.

ii. Huffman Coding

Hoffman coding was developed by David Huffman in 1951. Huffman coding is an entropy encoding algorithm used for lossless data compression. In this algorithm fixed length codes are replaced by variable length codes. When using variable-length code words, it is desirable to create a prefix code, avoiding the need for a separator to determine code word boundaries. Huffman Coding uses such prefix code [26].

iii. ARITHMETIC Coding Technique

Arithmetic coding can treat the whole symbols in a list or in a message as one unit [27]. Unlike Huffman coding, arithmetic coding doesn´t use a discrete number of bits for each. The number of bits used to encode each symbol varies according to the probability assigned to that symbol. Low probability symbols use many bit, high probability symbols use fewer bits [28]. The main idea behind Arithmetic coding is to assign each symbol an interval. Starting with the interval [0...1), each interval is divided in several subinterval, which its sizes are proportional to the current probability of the corresponding symbols [29]. The subinterval from the coded symbol is then taken as the interval for the next symbol. The output is the interval of the last symbol [30,31].

iv. SHANNON-FANO CODING

Claude E. Shannon was developed by Robert M. Fano. The algorithm evaluates the probability for each letter then assign code to each letter depending on its probability.shanon fanno algorithm gains its popularity during its simplicity. The algorithm is easy to understand and implement .the algorithm doesn't need high programming skills to implement it. In practical operation Shannon-Fano coding is not of larger importance. This is especially caused by the lower code efficiency in comparison to Huffman coding as demonstrated later.

Utilization of Shannon-Fano coding makes primarily sense if it is desired to apply a simple algorithm with high performance and minimum requirements for programming. An example is the compression method IMPLODE as specified e.g. in the ZIP format. The next algorithms show the pesudocode of Shannon fanno algorithm

Algorithm 2: Shannon fanno pesudocode

1. Create a probability table.
2. Sorting the table based on the probability and places the most frequent symbol at the top of a list.
3. Divided into equally two halves upper and lower which is having a same probability as much as possible.
4. The upper half of the list defined with „0" digit and the lower half with a „1".
5. Repeat the steps 3 and 4 for each of the two halves then further divide the groups and adding bits to the codes and stop the process when each symbol has a corresponding leaf on the tree.

v. Adaptive Huffman Coding

Adaptive Huffman coding is one of the most important and widely used algorithm it also called dynamic algorithm .the algorithm create code depending on the sequence of previous steps . The algorithm depends on Huffman coding algorithm, as the symbols are being transmitted that allows one pass encoding and adaptation to changing conditions in data. The benefits of one-pass procedure are that the source can be encoded in real time, though it becomes more sensitive to transmission errors, since just a single loss ruins the whole code. This algorithm can overcome the drawbacks of Huffman coding, because it doesn't depend on probability distribution to generate their code.

vi. Lempel-Ziv-Welch

The Lempel-Ziv-Welch (LZW) algorithm was created in 1984 by Terry Welch. It depend on internal memory called dictionary which be used to replace the characters this process enhances the compression operation [24]. The next algorithm show the pesudocode of the steps of lzw compression

Algorithm 3: LZW Encoding Algorithm

Step1: Initialize dictionary to contain entry for each byte. Initialize the encoded string with the first byte of the input stream.
Step 2: Read the next byte from the input stream.
Step 3: If the byte is an EOF go to step 6.
Step4: If concatenating the byte to the encoded string produces a string that is the dictionary: Concatenate the byte to the encoded string. go to step 2
Step 5: If concatenating the byte to the encoded string produces a string that is not in the dictionary: add the new string to the dictionary. Write the code for the encoded string to the output stream. Set the encoded string equal to the new byte. go to step 2.
Step 6: Write out code for encoded string and text.

IV. HYBRID CRYPTOGRAPHIC AND COMPRESSION TECHNIQUES

SubhamastanRao, Soujanya, Hemalatha and Revathi (2011), this researcher was developed a new approaches which can do the two operations at the same time.the algorithm is cost and time effective and less time consuming because the algorithm do compression and encryption at the same time, the encryption strengths of our methods are as good as any other encryption algorithms such as DES, triple DES, and RC5. In this approach the speed is very high because here we need to encrypt only the Huffman table rather encrypting the whole file which is to be transmitted. The encryption strength of the method of encrypting the Huffman table depends on the length of the encryption key [40]. John and Manimegalai (2012) classify the compression and cryptography into two categories the first one is independent category which treat each one as a whole and the second category interest by the interaction between the two techniques it is also called joint category. Independent encryption techniques can further be classified as heavy weight and light weight encryption algorithms. There are many algorithms available in the joint compression and encryption technique. Joint compression and encryption algorithms perform better in terms of speed and security when compared to independent encryption algorithms [41]. Ruchita and Swarnalata (2015), proposed that encryption and compression are done at the same time then it takes less processing time and more speed. They evaluate the performance with respect to different parameters. It shows basic information about cryptography and compression, & their techniques are applied on text file. For data security, combination of compression and cryptographic techniques are used. To secure our data more that's why we compressed the data first and then encrypt that compressed data [42].

V.CONCLUSIONS

The paper is presented some of compression and cryptography techniques which can be used in the data security operation .this survey opens the door for all researcher of security, especially data security in electronic commerce to develop a good model to enhance the security performance of e-commerce application.

REFERENCES

1. Al-Slamy, Nada MA. "E-Commerce security." IJCSNS 8.5 (2008): 340.
2. Chaffey, D. E-Business and E-Commerce Management, 2nd ed., Prentice Hall, Harlow, UK(2004).
3. Ackerman, Mark S., and Donald T. Davis Jr. "Privacy and security issues in e-commerce." New economy handbook (2003): 911-930.
4. Turban, Efraim, et al. "Electronic commerce: A managerial perspective 2002." Prentice Hall: ISBN 0 13.975285 (2002): 4.
5. Hou, J. , 'Research on database security of E-commerce based on hybrid encryption', Proceedings of the 2009 International Symposium on Web Information Systems and Applications, Nanchang, China(2009), pp. 363-366.
6. Oreku, George S., and Jianzhong Li. "Rethinking E-commerce security." Computational Intelligence for Modeling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on. Vol. 1. IEEE, 2005.
7. Yasin, Shazia, Khalid Haseeb, and Rashid Jalal Qureshi. "Cryptography based e-commerce security: a review." International Journal of Computer Science Issues 9.2 (2012): 132-137.
8. 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.
9. Bhanot, Rajdeep, and Rahul Hans. "A review and comparative analysis of various encryption algorithms." International Journal of Security and Its Applications 9.4 (2015): 289-306.
10. Schneier, Bruce. "Description of a new variable-length key, 64-bit block cipher (Blowfish)." International Workshop on Fast Software Encryption. Springer Berlin Heidelberg, 1993
11. Koko, Soheila Omer AL Faroog Mohammed, and Amin Babiker. "Comparison of Various Encryption Algorithms and Techniques for improving secured data Communication." IOSR Journal of Computer Engineering (IOSR-JCE) 17.1 (2015): 62-69.
12. Stallings, William. Data and computer communications. Pearson/Prentice Hall, 2007.
13. Kumar, SG Saravana, and A. Shanmugam. "Modified F–Function for Feistel Network in Blowfish Algorithm." gene (chosen randomly) 4.4 (2014).
14. Kahate, Atul. Cryptography and network security. Tata McGraw-Hill Education, 2013.
15. Rivest, Ronald L. "The RC5 encryption algorithm." International Workshop on Fast Software Encryption. Springer Berlin Heidelberg, 1994.
16. Islam, Md Nazrul, et al. "Effect of security increment to symmetric data encryption through AES methodology." Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2008. SNPD'08. Ninth ACIS International Conference on. IEEE, 2008.
17. Yasin, Shazia, Khalid Haseeb, and Rashid Jalal Qureshi. "Cryptography based e-commerce security: a review." International Journal of Computer Science Issues 9.2 (2012): 132-137.
18. Garrett, S. G. E., and P. J. Skevington. "An introduction to electronic commerce." BT Technology Journal 17.3 (1999): 11-16.
19. Rasmi, P. S., and Varghese Paul. "A Hybrid Crypto System based on a new CircleSymmetric key Algorithm and RSA with CRT Asymmetric key Algorithm for E-commerce Applications." International Conference on VLSI, Communication & Instrumentation. Kerala. Vol. 9. 2011.
20. Palanisamy, V., and A. Jeneba Mary. "Hybrid cryptography by the implementation of RSA and AES." International Journal of Current Research 33.4 (2011): 241-244.
21. Kuppuswamy, Prakash, and Saeed QY Al-Khalidi. "Implementation of security through simple symmetric key algorithm based on modulo 37." International Journal of Computers & Technology 3.2 (2012): 335-338.
22. Kuppuswamy, Prakash, and Saeed QY Al-Khalidi. "Securing E-Commerce Business Using Hybrid Combination Based on New Symmetric Key and RSA Algorithm." MIS REVIEW: An International Journal 20.1 (2014): 59-71.
23. Sidhu, Amandeep Singh, and M. Garg. "Research Paper on Text Data Compression Algorithm using Hybrid Approach." International Journal of Computer Science and Mobile Computing 3.12 (2014): 01-10.
24. Sharma, Neha, Jasmeet Kaur, and Navmeet Kaur. "A review on various Lossless text data compression techniques." International Journal of Engineering Sciences, Issue 2 (2014).
25. Nelson, Mark R. "LZW data compression." Dr. Dobb's Journal 14.10 (1989): 29-36.
26. Maan, Anmol Jyot. "Analysis and Comparison of Algorithms for Lossless Data Compression." International Journal of Information and Computation Technology 3.3 (2013): 139-146.
27. Shahbahrami, Asadollah, et al. "Evaluation of huffman and arithmetic algorithms for multimedia compression standards." arXiv preprint arXiv:1109.0216 (2011).
28. Witten, Ian H., Radford M. Neal, and John G. Cleary. "Arithmetic coding for data compression." Communications of the ACM 30.6 (1987): 520-540.

29. Kurre, Omeshwari, and Mr Amit Kolhe. "Video Compression with Wavelet Transform Using SPIHT and Arithmetic Coding Technique."
30. Sharma, Mamta. "Compression using Huffman coding." IJCSNS International Journal of Computer Science and Network Security 10.5 (2010): 133-141.
31. Li, Ze-Nian, Mark S. Drew, and Jiangchuan Liu. Fundamentals of multimedia. Upper Saddle River (NJ):: Pearson Prentice Hall, 2004.
32. Shannon, Claude Elwood. "A mathematical theory of communication." ACM SIGMOBILE Mobile Computing and Communications Review 5.1 (2001): 3-55.
33. Fano, Robert M. The transmission of information. Cambridge, Mass, USA: Massachusetts Institute of Technology, Research Laboratory of Electronics, 1949.
34. Ziv, Jacob, and Abraham Lempel. "Compression of individual sequences via variable-rate coding." IEEE transactions on Information Theory 24.5 (1978): 530-536.
35. Jain, Amit, and Kamaljit I. Lakhtaria. "A comparative study of lossless compression algorithm on text data." (2013).
36. Gallager, Robert. "Variations on a theme by Huffman." IEEE Transactions on Information Theory 24.6 (1978): 668-674.
37. Jani, Hardik, and Jeegar Trivedi. "A survey on different compression techniques algorithm for data compression." International Journal of Advanced Research in Computer Science & Technology 2.3 (2014): 364-368.
38. Vitter, Jeffrey Scott. "Design and analysis of dynamic Huffman codes." Journal of the ACM (JACM) 34.4 (1987): 825-845.
39. Ziv, Jacob, and Abraham Lempel. "A universal algorithm for sequential data compression." IEEE Transactions on information theory 23.3 (1977): 337-343.
40. SubhamastanRao, T., et al. "Simultaneous data compression and encryption." International Journal of Computer Science and Information Technologies 2.5 (2011): 2369-74.
41. Singh, K. John, and R. Manimegalai. "A survey on joint compression and encryption techniques for video data." Journal of Computer Science 8.5 (2012): 731.
42. Sharma, Ruchita, and Swarnalata Bollavarapu. "Data Security using Compression and Cryptography Techniques." International Journal of Computer Applications 117.14 (2015).
43. Xiaoqin, Lian, et al. "Application of the Advanced Encryption Standard and DM642 in the image transmission system." Computer Science & Education (ICCSE), 2012 7th International Conference on. IEEE, 2012.
44. Mathur, Milind, and Ayush Kesarwani. "Comparison between Des, 3des, Rc2, Rc6, Blowfish And Aes." Proceedings of National Conference on New Horizons in IT-NCNHIT. Vol. 3. 2013.

