International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 06 | June -2020 ISSN: 2582-3930

Hybrid technique of data Security by ECS (Encryption, Compression, Steganography)

Shishupal Sidar [1], Gourav Chandel [2], Deepak Garg[3]

[1,2,3]National Institute of Technology, Kurukshetra, Haryana, India

Abstract. The security of confidential data has long been a serious concern, and as a result, data security is in huge demand. Data is considered an important aspect of an asset and must be protected. Data Protection comprises data integrity, data authenticity, data confidentiality, and then some. Data is prone to possible security risks despite several measures for protecting data including steganography, encryption, and compression. The great reason is the disadvantages of these particular methods when certain factors are taken into account. Cryptography provides security, but it is expensive when time and space are considered. Compression methods encrypt data and save disk space, but attackers can easily decompress these compressed files. Steganography is a technique that hides data, does not encrypt it, and also has its advantages and disadvantages. This article details the approaches to data protection (compression, steganography, and compression) and their possible pros and cons. An extensive comparison between these techniques was generated and as a result, a hybrid approach called ECS (Encryption, Compression, and Steganography) was proposed. The proposed method utilizes particular approaches to overcome the limitations of existing techniques while also maintaining low complexity and minimal complexity and rich data security.

Keywords: Cryptography, Data Security, Decryption, ESC-(Encryption, Steganography, Compression).

1 Introduction

Data security is shielding data from threats. Any data transmitted over the Internet can be utilized for different purposes. Therefore, different strategies, for example, Compression, steganography, encryption cryptography decryption cryptography are utilized to shield Data from threats when they are shared over the Internet [1][2][3]. This works by blending plaintext into encrypted text and afterward once more. Keys are utilized for Data encryption: symmetric (Decryption and encryption are performed with a similar key) and Asymmetric key (decryption and encryption are performed utilizing public and private keys, separately) [1] [3]. Steganography is a strategy for concealing a message in a veil (usually an image) with the goal that assailants can be certain that no private data is accessible. This shrouds the way that a secret message is being transmitted [2]. The way toward decreasing the number of bits involved by the encrypted text is called Compression. This works by wiping out Data redundancy and Data deduplication. This gives lower memory utilization, minimal, and requires minimum time to move a record over the network [1].

2 Challenges:

(A) Disadvantages of individual approaches:

- **Encryption**: It is suspicious because the data is being converted to a meaningless form.
- **Compression**: This is a method of data compression, and algorithms are weak to ensure data security.
- **Steganography**: not enough to guarantee data security, because after detecting the storage medium, the hidden data can be easily decrypted.

(B) To determine the best possible approaches while maintaining high data security and lower computational speed/time complexity:

The main task in developing a hybrid security system is to choose the best possible encryption, compression, and steganography approaches to ensure high data security and low time complexity.

(C) Security versus time complexity:

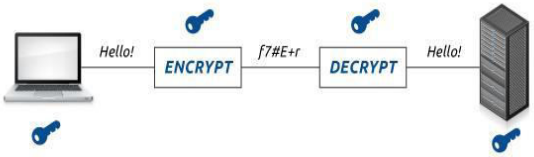
To keep up a high level of security, which is the most significant necessity in this situation, the complexity of the speed/time count is additionally diminished. To ensure the system staggered security system is required.

3 Cryptography

Data security, cryptography plays an important role. Crypto means something hidden. This is the study and use of secure correspondence methods in the field of vision of unapproved parties. It is dedicated to the development and analysis of protocols that do not allow malicious third parties to receive data exchanged between two objects, thus aspects of Data security [1][3].
Following are the classified categories of Cryptography:

- Hash functions
- Symmetric key cryptography and,
- Asymmetric key Cryptography

3.1 Symmetric-key: These are cryptographic algorithms that utilize the equivalent cryptographic keys to encrypt plain text and decrypt encrypted text. The keys can be indistinguishable, or a basic change called a common key or shared secret encryption, can be performed between two keys. With symmetric encryption, a single key is utilized to



encrypt and decrypt traffic [3].

Figure 1: Symmetric key, a single key for encrypt and decrypt.

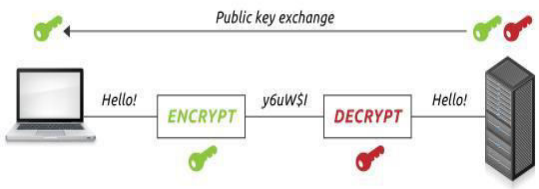
Advanced Encryption Standard (AES): One of the most widely utilized symmetric encryption algorithms, it is evaluated to be at least six times faster in comparison with DES. The key size used in DES was too small and that led to the development of a replacement. With increasing computing power in recent times, it was found prone to possible key search attacks. As an alternative, a triple DES algorithm was proposed which was then dropped due to its slow computation speed [12].

Here are some of the key features of the Advanced Encryption Standard:

- Offers detailed design and specification.
- Symmetric Block cipher and key.
- 128/192/256 bit keys.
- Faster and more robust than Triple DES

Data Encryption Standard (DES): DES, expanded as Data Encryption Standard works on block encryption technique with the help of symmetric key circulated by NIST (National Institute of Standards and Technology) [13]. The DES encryption technology is based on the 16-round structure encryption with a block size of 64 bits. Even though the length of the key is 64 bits, the corresponding DES key length is 56 bits due to the 8 non-significant bits (these 8 bits only offer their role as control bits.)

3.2 Asymmetric key: Asymmetric Keys are different from symmetric keys, they work on the principle of two distinct keys, a private key that is kept confidential to the owner during the communication, and a public that can be distributed widely among users. The generation of these key pairs depends heavily on the mathematics-based cryptographic algorithms to obtain unidirectional functions. This encryption technique uses different keys during the



decryption and encryption process, therefore, any user can encrypt data using the public key available encryption key. [3].
Figure 2: Asymmetric encryption,(public-key exchange)
Asymmetric types:

- RSA

- ECC
- ElGamal

(i) **RSA (Rivest Shamir Adelman):** RSA is an asymmetric cryptographic algorithm that is used to encrypt and decrypt data. During the process, the encryption key is shared publicly while the key to decrypt data is different from the encryption key. [15].
RSA is an asymmetric algorithm. The main application of the RSA are:
1. Key exchange or authentication using encryption of small data (mostly symmetric keys)
2. Digital signatures.

(ii) **Elliptic Curve Cryptography (ECC):** ECC is a public-key encryption algorithm technique defined based on the algebraic structure of elliptic curves spanning finite fields. It requires a comparatively smaller key than non-ECC encryption to provide a similar level of protection. (256-bit ECC protection has equivalent protection achieved by RSA 3072-bit crypto).

(iii) ElGamal encryption: ElGamal is a public-key based cryptographic algorithm that uses an asymmetric key-based technique for the data encryption and communication between two parties. It is based on the complexity of finding a discrete logarithm in a cyclic group.

Table1: RSA vs ECC vs ElGamal [4]

Comparison Factor	ECC	ElGamal	RSA
Security	the theory of elliptic capacities	the trouble of computing discrete logarithms.	the trouble of calculating big integers.
Base	Elliptic Functions	Logical functions	Big Number
Encrypt size keys-256	7098	8242	37
Decryption size keys-256	37	3932	37
generation time size keys 256	895	6451	1957

3.3 Hashing functions: It is a mathematical function that is used for the conversion of one input value into another compressed numeric value. The input fed into the Hash function can be of variable length but the output always turns out to be of fixed length. This approach claims full responsibility for the message integrity since the hash value on both receiver's end and the sender's end should be the same.[3]

Table 2: Comparison & Conclusion on Cryptography

Feature	Symmetric	Asymmetric	Hash function
Keys required	1	2	0
Approved key length by NIST	128 bits	2048 bits	256 bits
Often used	AES	RSA	SHA
key compromise effect	Senderlose and receiver lose	Loss only for asymmetric key	N/A
Speed	Fast	Slow	Fast
Complexity	Medium	High	Medium
Examples	AES, DES	RSA, Elgamal, ECC	Sha 224, Sha 256,

Table 3: Md5 VS SHA1

Comparison Factor	SHA 1	Md5
Security	Moderate	Poor
Output size	160 bits	128 bits
find two messages provide a similar output	2^{80} operations (4 rounds * 20 steps)	264 operations (4 rounds * 16 steps)
find original message correspond to the output	2^{160} operations	2^{128} operations
Speed	Slower	Faster
Mathematical expression	a= (e + Process P +S5 (a) + W [t] + K[t]), b= a, c= s 30 (b), d=c, e= d	a = b + ((a + Process p (b, c, d) +M[i] +T[k]) <<<s), b=b, c=c d=d
Attacks	SHAttered attack	Collision attack

4 Compression:

Compression decreases the size of data (audio, image, text, video, etc.). This is an approach to diminish the size of source data utilizing certain encryption techniques, which, thusly, spare memory, accelerate document moves, lessen capacity expenses, and network data transfer capacity, just as decrease size of the file and spare disk space.

Lossless Compression: With lossless compression, we get the same data as before compression. This is the only lossless reason. Compression is applied to text files since no loss is allowed in text files.

Lossy Compression: When compressing at a lossy, we don't get the same data as before compression, and there is some data on the losses. Therefore, these types of compression methods apply to images, audio, and video, because some losses can be allowed here [8]

Table 4: Comparison of Compression techniques

Comparison Factor	Huffman	LZW	Run Length
Speed	Fast execute	Fast Compression	Fast execute
Drawbacks	Problematic due to different code lengths. [10]	The table ranks are troublesome.	A high compression ratio cannot be achieved.
Time Complexity	O(n(log n)	O(n)	O(n)
Space Complexity	O(k) for the tree and O(n)	O(n)	0(2n)
Decoding (Input)	Required.	No prior data	No data required.
Cost	Additional costs transmit.	Low cost The transfer cost is less.	Sophisticated decoding process.
Applications	JPEG, GIF, MPEG, ARJ	PDF,TIFF,GIF	BMP,TIFF,BMP

5 Steganography:

Steganography is a technique that is utilized in communication and permits you to conceal secret data in any medium. For this situation, the sender gets certain data, encoding it in different ways. The medium used to shroud data can be image, text, video, or any audio the message is installed on the spread utilizing an algorithm and steno record and is sent through the communication channel.

Table 5: Comparison of Steganography Methods

Comparison Factor	Images	Audio	Text
Cover media	Image	Audio	Text
More prone attacks	Less	Least	Most
Difficult	Less	difficult	Least
Main Feature	The most famous medium due to its high level.	The audio on the cover as "noise" and frequency that is out of earshot.	Used a common compound language to hide a cryptic message.
Applications	IP packets/ TCP.	Phone Calls, Skype.	CSS, emails, SMS texting,
Undetectability	Very large	Poor	very small
Bit Rate	Good	Medium[13]	Poor
Resistance modification	Very good	Weak	Average

6 Literature Review

Uses two methods for the encryption of data. The first method follows the concepts of substitution and transposition cipher. Ambiguity arises in the case of shifting letters of the input text and diffusion by swapping and reversing the input text. Whereas the second method uses a different algorithm where letters at even position are shifted by n letters, and letters at the odd position are decremented by n letters [1].brings light to the fact that the performance of algorithms like RSA, ElGamal, and Elliptic Curve reduces with an increase in key size. As the size of the input text increases, key generation, encryption, and decryption increases proportionally as well. The author also enlightens the fact that encrypted messages usually get larger than the original message that leads to an excess charge in sending an encrypted message. With their future, the author promises to introduce some kind of compression over the encrypted texts that can significantly reduce the overhead in sending encrypted messages [4]. This paper tries to preprocess the data before it is masked. In the preprocessing phase, the data is compressed, encrypted, and digitally watermarked. The compression reduces the size of the data so that the number of messages to be embedded into the same cover image increases. Encryption happens over the plain text with a key which alters it into encrypted text, and the final watermarking phase ensures sender authentication, data verification, and file reconstruction [5].
Uses Filter Bank Cipher over Galois Field for the encryption of plain input text, implements DWT based steganography on the encrypted text, uses embedding data, Stegano image, and encrypted message generation to achieve the final encryption. The text is then finally decrypted at the receiver's end using the Synthesis filter bank and perfect reconstruction lifting scheme [7]. Uses two-step encryption where the text is first encrypted using the AES algorithm whereas the resultant cipher text is encrypted again using the ECC algorithm. The final encrypted text is then compressed using the LZW compression technique and embedded with additional data [8]. Proposes encryption of data happens before data transmission to ensure data security. The Hybrid Encryption Model is utilized for the encryption of data. The encryption is finished by data cleaning for data normalization and data de-

duplications. Duplication Data detection technology first isolates the data document into a gathering of data squares, assesses fingerprints for each data square, at that point utilizes fingerprints as keywords for playing out a Hash search. The encryption model uses ECC encryption for encoding the Hash table and the Symmetric encryption method to re-encryption the enormous data handled after the primary period of encryption [10].

7 Proposed Model

RSA works extraordinarily well with Asymmetric encryption, and can likewise be joined with symmetric encryption keys (modify Playfair) to make a system that guarantees data security with minimal complexity in time and space. Compression strategies help pack repetitive data bits, and steganography helps hide data in any type of multimedia, so steganography and Compression techniques can likewise frame a safe half breed System. Steganography and compression are utilized consecutively to encrypt plain text to guarantee a significant level of data security. This can be used to protect the data using the above methods in combination with each other along with their algorithms. We use encryption approaches, then steganography approaches and compression approaches sequentially to encrypt plain text to ensure the high level of data security shown in Figure 3.

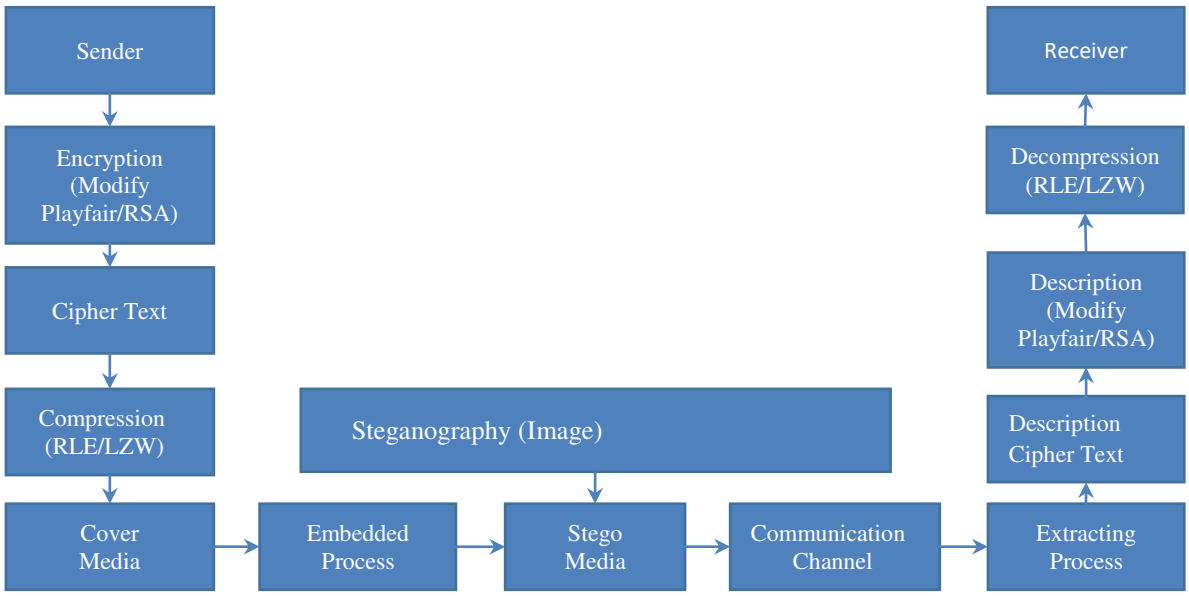


Figure 3: Proposed Model of ECS (Encryption Compression Steganography)

7.1 Example:

Stage 1:

This technique is a blend of the symmetric cipher (Modify Playfair) and Asymmetric cipher (RSA). Since symmetric encryption experiences a key trade issue, this is wiped out by blending it in with Asymmetric encryption.

Make a 7*7 matrix with a keyword. Fill in the rest of the sections in the matrix utilizing the rest of the letters in alphabets, maintaining a strategic distance from duplicates. Encryption plain text utilizing the array produced by the modified Playfair encryption rules.[15]

This example takes any of the four edges as a beginning stage and devours the matrix in a concentric spiral.

A SPIRAL PATTERN USING 7 * 7 MATRIX

1	2	3	4	5	6	7
24	25	26	27	28	29	8
23	40	41	42	43	30	9
22	39	48	49	44	31	10
21	38	47	46	45	32	11

20	37	36	35	34	33	12
19	18	17	16	15	14	13

Consider a keyword K = “ace”. K contains 3 characters.
Primary state and utilizing '!' as filler character and ' ~ ' as a padding character

PRIMARY STATE

Char.	a	b	c	d	e	…….	!	~
index	0	1	2	3	4	…….	47	48

In the first place, the list of K is determined. In this way, index (K) = [0, 2, 4].
Let, A is empty. After affixing characters from K,

A = ['a', 'c', 'e'].

Also, re-listed, which is appeared in the table.

AFTER A REARRANGE AND RE-INDEXING OPERATION

Char.	b	d	f	g	h	…….	!	~
index	0	1	2	3	4	…….	47	48

At that point, a square of characters ['b', 'f', and ‘h’] is extricated from utilizing index (K). Characters are attached to the array.
Presently, the array, A = [a, c, e, b, f, h]. Along these lines, all the characters are separated. Finally, A = ['a', 'c', 'e', 'b', 'f', 'h', 'd', 'i','k', 'g', 'l', 'n', 'j', 'o', 'q', 'm', 'r', 't', 'p', 'u', 'w', 's', 'x', 'z', 'v', '0', '2', 'y', '3', '5', '1', '6', '8', '4', '9', ')', '7', '$', '+', '(', ',', ':', '&', ';', '', '/', '!', '=', '~']

Now, applying a spiral pattern (heading: clockwise, beginning stage: upper-left edge) to the data in A, we get the matrix that appeared in table 5.

THE KEYWORD "ace" IN SPIRAL PATTERN

a	**c**	**e**	**b**	**f**	**h**	**d**
z	**v**	**0**	**2**	**y**	**3**	**I**
x	**(**	**,**	**:**	**&**	**5**	**K**
s	**+**	**=**	**~**	**;**	**1**	**g**
w	**$**	**!**	**/**		**6**	**l**
u	**7**	**)**	**9**	**4**	**8**	**n**
19	**18**	**17**	**16**	**15**	**14**	**j**

Encryption: Plaintext “bb″

Plaintext processing:
b! → use filler character
b~ → use padding character
Block substitution:
b! → e/
b~ → 2/

Decryption: Cipher text “e/2/”

Block substitution:
e/ → b!
2/ → b~

Omitting padding and filler character and ignoring filler and padding characters, retrieved plaintext is “bb”.

Then RSA applies here the subsequent text will be in plain text, which will in the long run be encrypted with the RSA public key.
Stage 2: Hide the encryption acquired with the image steganography.
Stage 3: Apply any of the compression strategies (LZW/RLE) to lessen the additional bits and hence get a Compression file.

8 Conclusion

Security is just a myth. With new threats against data rising every day, protection techniques such as cryptography, steganography, compression, and other renowned techniques cannot protect the data alone [8]. At any point in time, the encrypted data can be deciphered by a cryptanalyst. Steganography only serves the purpose of masking plain text behind some digital format. The mask is bypassed; every data can be accessed in its plain form. However, if a text is only encrypted, it may raise suspicion and can be easily compromised after successful decryption. To churn out a better solution to the existing problem, this paper proposes a collective approach of all the mentioned techniques. For instance, RSA hybridization (as Asymmetric technique), modify player (as a symmetric methodology), image steganography, and consecutive length encoding, or LZW (as Compression) is a decent mix to guarantee high security of data and low complexity. Future work may incorporate different blends and the investigation of cryptography (modern and secure), cryptography and steganography, or cryptography in the mix with steganography, and will keep on being bundled to guarantee security and lessen multifaceted nature of presence.

Reference:

[1] Kaushal A., Enhancement in Data Security using Cryptography and Compression. In: International Conference on Communication Systems and Network Technologies, pp 212-215(2017).

[2] Vaithiyanathan, A Survey on Image Steganography IEEE International Conference on Technological Advancements in Power and Energy (TAP Energy) (2017).

[3] Rajani.T., Importance of Cryptography in Network Security. International Conference on Communication Systems and Network Technologies (2013), pp 462-467(2013).

[4] Seral D., Sms Security: An Asymmetric Encryption Approach. In Sixth International Conference on Wireless and Mobile Communications, pp 448-452(2010)

[5] Gaba, J., Sharma, M.k.: A Review Based Study of Hybrid Security Schemes Based on Compression, Encryption, and Steganography. In: International Journal of Engineering Trends and Technology, vol. 4(7), pp. 3243-3246 (2013)

[6] Ibrahim, A.M.A., Mustafa, M.E. Comparison Between (RLE And Huffman) Algorithms for Lossless Data Compression. In: International Journal of Innovative Technology and Research, vol. 3(1), pp. 1808-1812(2015)

[7]En.wikipedia.org.(2019). RSA_(cryptosystem)[online] Available_at:https://en.wikipedia.org/wiki/RSA_(cryptosystem) [Accessed 25 Mar. 2019].

[8] Search Storage. (2019). What is data compression? - Definition from WhatIs.com. [online] Available at: https://searchstorage.techtarget.com/definition/compression [Accessed 25 Mar. 2019].

[9]Anon,(2019).[online]Available_at:https://www.researchgate.net/publication/273011398_A_Secure_Data_Communication_System_Using_Cryptogaphy_And_Steganography [Accessed 24 Mar. 2019].

[10] Manjula Y., Enhanced secure image steganography using double encryption algorithms 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)

[11] Hui Y Research on Real-time Analysis and Hybrid Encryption of Big Data 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD).

[12]En.wikipedia.org.(2019).Data_Encryption_Standard[online]Available_at:https://en.wikipedia.org/wiki/Data_Encryption_Standard [Accessed 25 Mar. 2019].

[13] Maan, A.J.: Analysis and Comparison of Algorithms for Lossless Data Compression. In: International Journal of Information and Computation Technology, vol. 3(3), pp. 139-146(2013)

[14] En.wikibooks.org. (2019). Steganography/Covers - Wikibooks, open books for an open world. [online] Available at: https://en.wikibooks.org/wiki/Steganography/Covers#Audio [Accessed 25 Mar. 2019].

[15] AhnaftahmidShakilMd.and. Islam Md," An Efficient Modification to Playfair Cipher". ulab journal of science and engineering vol. 5, no. 1, November 2014 (ISSN: 2079-4398)