ISSN 2278-3091
Volume 9 No.2, March -April 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse226922020.pdf
https://doi.org/10.30534/ijatcse/2020/226922020

Data Hiding using Meaningful Encryption Algorithm to Enhance Data Security

Devishree Naidu[1], Shubhangi Tirpude[2], Kanak Kalyani[3], Vrushali Bongirwar[4], Tejasvee Sharma[5]
[1]Assistant Professor,ShriRamdeobaba College of Engineering & Management, Nagpur, Maharashtra, India, naidud@rknec.edu
[2] Assistant Professor, ShriRamdeobaba College of Engineering & Management, Nagpur, tirpudes@rknec.edu
[3] Assistant Professor, ShriRamdeobaba College of Engineering & Management, Nagpur, kalyanik@rknec.edu
[4] Assistant Professor,ShriRamdeobaba College of Engineering & Management, Nagpur, bongirwarvk@rknec.edu
[5] Treasury Transaction Technology, Barclays Global Services Pvt. Ltd.,Pune, Maharashtra, tejasvee.sharma@barclays.com

ABSTRACT

Encryption allows you to securely ensure confidentiality. The secrecy of embedded data is essential in the field of data security and privacy. Hence, the data should be hidden from the ones who intend to corrupt it or use it in unauthorized manner. That's why it is important to enhance the secrecy of the encrypted data, hide the existence of confidential data and ensure data security while device to device data transfers. Only the owner should have the access to the information. If we are ever being watched, inadvertently or not, here we hide our data by applying symmetric encryption in such a manner that encrypted data is often represented in non-understandable language and easily identified as an encrypted form. The scheme makes use of global database system for information hiding and also overcome limitation of Steganography . This paper discuss about enhancement of communication security by embedding secret messages into an inconspicuous carrier and thereby transmitting them to receivers.

Key words: Decryption**,** Encryption**,**Security,Steganography

1. INTRODUCTION

Data plays vital role in fundamental functioning and structuring of an organization or on individual level. In Age of Information, data security takes topmost priority because its compromise could result in unauthorized access, cyber-attacks and espionage.

To maintain confidentiality of the data, it needs to be encrypted to keep data secure. This Scheme aims to implement the concept of Cryptography by introducing the language of decryption as meaningful paragraphs with domain distinguished than the original content to deceive the unauthorized access from interpreting the information.

Data or information is very crucial to any organization or any individual person. Thus securing the information becomes all the more necessary. The information can be transmitted without permission and modified, used and misinterpreted to represent an individual or maybe used to attack. If information reaches the unauthorized person they might arise a lot of complications. Hence there is a need to hide the data so that a third person or irrelevant person cannot extract the exact message. As the growth and development of technology is increasing, the concern for the safety and security of data is increasing equally. We need to share the data in encrypted form on open communication channels to ensure its security. The norm of cryptography mainly focuses on the methodology with which a message is transformed into a covert form and then shared over public communication channels in order to maintain the confidentiality of the message.

2. LITERATURE REVIEW

S. Srivastava and N. Gupta proposed a new technique on extended play cipher in advent with design of 8*8 matrixes. The encryption technique takes care of handling spaces between words and even numerical, special character too. The cryptanalysis task were performed through launching of avalanche effect and test shown by the implementation firmly follows that breaking of cipher in case of known plain text attack is impossible. The scheme makes use of LFSR technique that showed output sequences differs by 48 bit

which is quite substantial and enough to declare strong cipher against any cryptanalytic attacks. This method fails to work against high bandwidth and high storage capacity[1].

AmjadHussainZahid in his proposal presented a novel technique for designing cryptographically strong substitution-boxes using cubic polynomial mapping. The cubic polynomial mapping was quite proficient to map the input sequence to a strong 8 × 8 S-box meeting the requirements of a bijective function. The design of simple S-BOX substitution were tested on real-time use case which was evaluated analytically using standard performance criteria including nonlinearity, bijection, bit independence, strict avalanche effect, linear approximation probability, and differential uniformity. The performance results were equated with mostly scrutinized S-boxes to ascertain Cryptographic forte. The critical analysis endorse showed the proposed S-box construction technique considerably innovative and effective to generate cryptographic strong substitution-boxes[3].

Mohammad Ali BaniYounes and AmanJantan used encryption scheme to securely transmit data in open networks and to protect confidential image data from unauthorized access. They focused on multimedia data such as images for which they used a block-based transformation algorithm based on the combination of image transformation and a well known encryption and decryption algorithm called Blowfish. The original image was divided into blocks, rearranged into a transformed image using a transformation algorithm and then the transformed image was encrypted using the Blowfish algorithm. They improved lower correlation and higher entropy with increasing number of blocks of smaller block size [4].

Nishtha Mathura paper proposes an extension of a public-key cryptosystem to support a private key cryptosystem which is a combination of Advanced Encryption Standard and ECC. A hybrid encryption scheme has been used to increases competency and minimizes its drawbacks. The parameters were mainly focuses on the key length, no. of iterations and the type of side channel attack in their implementation. The key length was shown to be increased to 192 bit and the no. of iterations were noted as 12 that has shown improvement in performance to secure the content [8].

Mohammad Reza Najaf Torkaman describes DNA Steganography techniques to conceal secret session key which is transferred among sender and receiver throughout unsecured channel. They developed a protocol in which attackers are not aware of exchanging session key which is hidden by DNA data hiding method. They used DNA Steganography algorithm which uses DNA reference sequences selected from EBI database. A work on approximately 163 million DNA sequences from EBI database were tested which was impossible for attackers to find whether or not there are secret message hidden in DNA sequences. Even if attackers know a secret message is embedded in DNA sequences, it is impossible to guess the correct sequence among 163 million DNA sequences [9].

3. PROPOSED APPROACH

The scheme aims to encrypt any confidential information to another meaningful text which does not convey its true meaning and thus maintaining the authenticity of the information without compromising its security. The meaningful encrypted text is in the form of a report of specific domain. The domain of the report depends on the key being used in this cipher. This key is a set of databases containing attributes used for generating the report. These attributes are numbered starting from zero in the databases. Two prominent part of algorithm are first the database for target data mapping and data to be encrypted that is given as input to the system, another is a Key Generation, here the database used for mapping that is shuffled randomly using Fisher-Yates Shuffle algorithm and is encrypted using any standard encryption algorithm such as RSA, DES. After these two processes, database is ready to be implemented for Cryptography.

3.1 Meaningful Encryption Algorithm

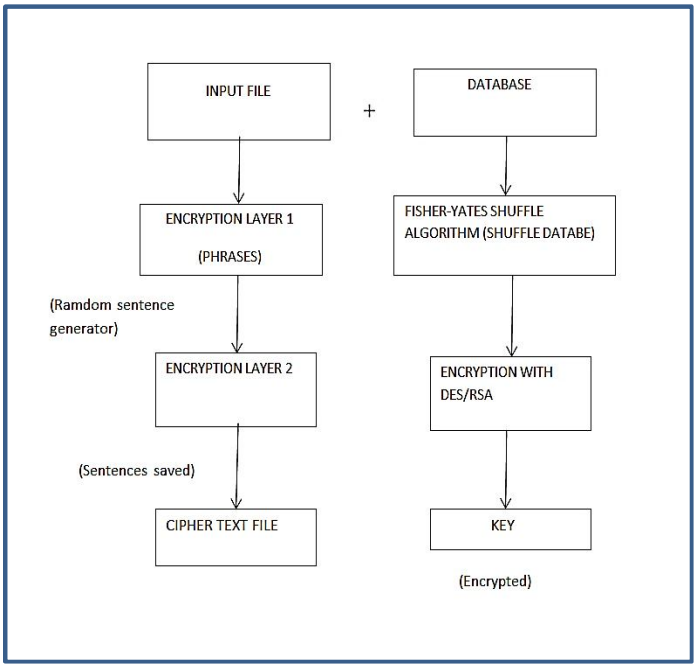

Figure 1: Work Flow of meaningful Encryption Algorithm

Algorithm :

1. System request to input a file for encryption, In parallel to this Database is uploaded for the relative context to map the contents in the form of report.

2. The database now shuffled using standard algorithm Fisher-Yates Shuffle for randomness.

3. These database get encrypted using RSA, including key generation and encryption ,to compose final database

4. Now in parallel Input file will follow the process of layer 1 encryption where words formations takes place through mapping of attributes value without any meaning.

5. Layer 2 encryption is followed where different types of grammar are applied to make the sentences meaningful. This step makes use of random sequence generator to make file as complex as possible. The key history of database is also stored in encrypted form at this layer.

6. Now meaningful sentenced mapped to the typical database content is the final encrypted version that is called as meaningful encrypted file as shown in Figure 1.Final output file.

7. The decryption steps are followed in reverse order to retrieve original information.

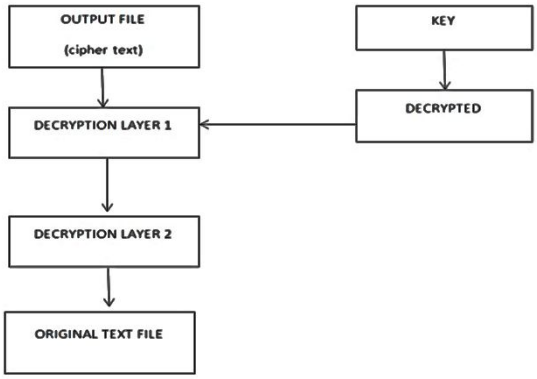

Figure 2: WorkFlow of Decryption Algorithm

3.2 Layer 1 and layer 2 Sub Encryption Process

The encryption process considers the input as any plain text. As an example for security concern some weather information selected to be kept confidential. Further encryption process makes use of different database to generate some meaningful composition as encrypted data. Ciphertext is generated for a Weather report of different locations that requires one database contains location names and the other contains weather specific attributes (Both databases are shuffled).

0	Amritsar
1	Delhi
2	Mumbai
3	Nagpur
4	Bangalore
…	…

Table 1: Location Information

0	Humidity
1	Cloudy
2	Sunny
3	Rainy
4	Foggy
…	…

Table 2: Weather Information

The Algorithm steps for Layer1 and layer 2 as follow

1. Process starts with pairing the two words.
2. Now select first letter of second word note its ASCII value
3. Calculate modulo 97 and note the result
4. Now check the numeric result value in location table get it to the output
5. Group first two letters with any letter of second word. Note their individuals ascii value and perform modulo 97
6. Now interpret first ASCII value that is chosen for attribute name. The second and third ASCII value will be interpreted as value at that attribute.
7. Continue the above procedures i.e. pair 1 group1-mapping till the message ends.
8. The ungrouped letters are assigned to select attribute for tagging at the end report.
9. Similarly for phrases handling, random sequence generators are used in the encrypted layer to generate meaningful text with sentences.

Sample Test Case:

If we consider any plain text as “hello world” **then we start with p**airing the words in the plain text

Pair 1- “hello world”

i. first letter of second word - “w”

ASCII value of w - 119

- 119 % 97 = 22

ii. The value corresponding to 22 in location database is Kasauli.

iii. Group the first two letters of "hello" with one letter of "world"

iv. Group 1- "h e o"

i. ASCII value of "h" will give the attribute name and ASCII value of "e" and "o" will give the value of this attribute.

ii. ASCII Value of "h"- 104

iii. - 104 % 97 = 7

iv. ASCII Value of "e"- 4%97 = 4

v. ASCII Value of "o"- 111%97=14

vi. The value corresponding to 7 in the weather attribute database is Thunderstorm and the value for this attribute will be 4 - 14

vii. Same procedure will be followed for the remaining groups of letters

A sample report on observations based on the above calculation at layer-2 information is as follow:

Survey Id: #902

We have Observed following at Kasuali:

Indexes for Thunderstorm are: 4-14

For Lightening indexes reported are: 11-17

Satellite Code: 14.-11.3.

Figure 3: Final Output Encrypted version

The Algorithm is Symmetric in Nature. Therefore, a systematic reverse approach is used on Encrypted Cipher file with encrypted key and to retrieve back file with original contents. Figure 1 and Figure 2 highlight the work flow.

4. PERFORMANCE ANALYSIS

All file text formats are accepted successfully and found flexible while operation on data. As described in table 3.we analyzed with different data sizes and recorded time for encryption to check for accuracy as the performance depends on the significant key parameters like size of the data, size of the database, encryption technique used for the key generation and attack mechanism to find vulnerabilities in the techniques.

Sr. no	Input file	File size	Encryption time in (sec)
1.	Tiny Text file	209 bytes	0.00369
2.	Text File of few sentences	4.17 kb	0.011
3.	PDF File of few sentences	587 kb	2.94
4.	Text File with bulky sentences	7.64 Mb	10.14s

Table 3: Analysis File Encryption

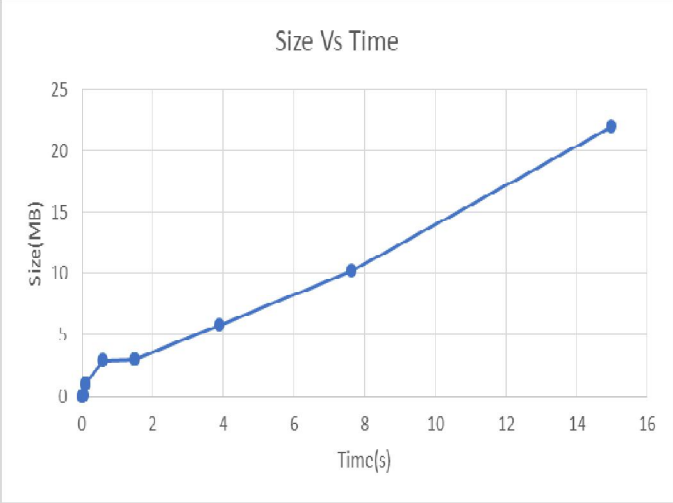

Figure 4: Performance Graph analysis for Encryption

4.1. Security Database

Database is encrypted and validated to avoid the access and modification by unauthorized sources, more improved could be done into database we could make it more secure by also applying a compression algorithm after the encryption to ensure better security.

4.2. Randomization of Mapping

The Mapping of the pairs with their database equivalent targets is the essential factors if overall degree of randomization will be less, security of the algorithm will go down. With introduction of Randomization not only it enhances the security but also introduces randomness in the mapping causing the similar inputs to generate distinguished

outputs. More complex and permutation-round base randomization of the database could ensure more complex network of mapping but at that same time would increase challenges for successful decryption.

4.3. Client-Server Model Configuration

Two authorized Client who are exchanging the secret information via the algorithm mentioned above are connected through an authenticated server. Server sends them some default settings as what database to use and how it is shuffled and key details. Two clients can send data to each other using these settings which can be changed dynamically to enhance the confidentiality and integrity of the data to be exchanged, this feature could ensure enforced and continuous safety of the data.

4.4. Accuracy

The accuracy of the system remain 100% as long as neither key nor encrypted data is altered or modified which could result in less accuracy and lost in data in certain portions of the data, but the entire encrypted file even though is independent modules but is not vulnerable to as each independent modules is further branched into groups and their encryption is based on the these groups conglomeration along with the randomized database as the keys are applied to avoid patterns.

4.5. Cases of Cryptanalysis Failure:

A. Absence of Database: Without a database which is our primary key, crypt-analysis cannot crack the code irrespective of permutations attempted.

B. In case of dynamism where with different database and new attribute names along with format permutation affecting changes in final encryption number of combination and cracking techniques also changes.

4.5. Usefulness of Meaningful Encryption

The algorithm has various applications in distinguished domains of our societies:

- Defense and Any Government Ministry:
  They are required to hold onto sensitive and confidential information and also to avoid it falling into wrong hands, a meaning encryption not only safeguards the information but also helps in avoiding the doubt of text holding any crucial information
- IT Industry and Service Infrastructure:
  Any domain from IT industry to everyday where we need to hide our important and crucial pieces of information from the eyes of Data Thieves who not only steals them but misuses them, in such cases a meaningful based encryption not only will protect your data but will also prevent data theft attacks as a meaningful text with no relevance to original data will avoid the suspicion of text holding any valuable information.
- Surveillance and Sensor Network:
  The algorithm could also prove beneficial for Surveillance and Sensor network as chip on scans and keeping the information stored into other target domain language.
- Settings based Client Server:
  Two authorized Client who are exchanging the secret information via the algorithm mentioned above are connected through an authenticated server. Server sends them some default settings as what database to use and how it is shuffled and key details. Two clients can send data to each other using these settings which can be changed dynamically to enhance the confidentiality and integrity of the data to be exchanged

5. CONCLUSION

The use of random sequence generator to convert sentences and introduction of layer 1 and layer 2 module for meaningful encryption makes the data more secure to existing encryption systems and proposes new features to hide the existence of confidential data and thereby enhancing the secrecy of the encrypted data. The benefits of this cipher can be summarized as; it allows secret transmission of messages without the fact that transmissive nature is discoverable. Double security of data is ensured through additional layer to standard encryption. Use of multiple databases of different domains creates the task of crypt-analyst an unauthorized user more difficult.

This cipher is thus beneficial for many business and military applications which include secret transmission of data. It encrypts any confidential information to another meaningful text which does not convey its true meaning and thus maintaining the authenticity of the information without compromising its security.

More improvements can be done to provide conversion for image information and different file formats supported on the cloud. Additional features can be envisioned towards modification of original information and strong authentication scheme may be build for strong system.

REFERENCES

[1] S. S. Srivastava and N. Gupta, “A Novel Approach to Security using Extended Playfair Cipher”, International Journal of Computer Applications (0975 – 8887), vol. 20, no. 6, (2011) April.

[2] Abd El-Latif AA, Li L, Zhang T, Wang N, Song X, Niu X (2012) Digital image encryption scheme based on multiple chaotic systems. Sens Imaging 13(2):67–88CrossRef [Google Scholar].

[3] Amjad Hussain Zahid, Muhammad Junaid Arshad, “An Innovative Design of Substitution-Boxes using Cubic Polynomial Mapping”, Symmetry 2019, 11(3), 437; https://doi.org/10.3390/sym11030437.

[4] Bani-Younes MA and Jantan A (2008) Image encrytion using block based transformation algorithm. Int J ComputSci (IAENG), 407– 415[Google Scholar].

[5] Bansal R, Chawla R and Gupta S (2016) A comparison of image encryption techniques based on chaotic maps, accepted in INDIACom - IEEE Conf [Google Scholar].

[6] Devishree Naidu, ShubhangiTirpude and VrushaliBongirwar (2019) “ Novel Idea of UniqueKey Generation and Distribution using Threshold science to Enhance Cloud Security” Smart Trends in Computing and Communications, Smart Innovation, Systems and Technologies 165, Proceeding of Smartcom 2019, https://doi.org/10.1007/978-981-15-0077-0_39

[7] RSA Laboratories,http://www.rsa.com/ rsalabs/node.asp?id=2193.

[8] Rahman M etal. “Development of Cryptography-Based ecure Messaging System”, Telecommunication System Manage 2016, Vol 5(3): 142DOI: 10.4172/2167-09-19.100014.

[9] Nishtha Mathura, Rajesh Bansode,7th International conference on Communication, Computing and Virtualization 2016 AES Based Text Encryption Using 12 Rounds with Dynamic Key Selection, https://doi.org/10.1016/j.procs.2016.03.131.

[10] Mohammad Reza Najaf Torkaman1, “Innovative Approach to Improve Hybrid Cryptography by Using DNA Steganography’ International Journal on New Computer Architectures and Their Applications (IJNCAA) 2(1): 225-236 The Society of Digital Information and Wireless Communications, 2012 (ISSN: 2220-9085) 225.

[11] Lauridsen, M.M.; Rechberger, C.; Knudsen, L.R. Design and Analysis of Symmetric Primitive; Kgs. Lyngby, Technical University of Denmark: KongensLyngby, Denmark, 2016. [Google Scholar].

[12] Mohamed, K.; Nazran, M.; Pauzi, M.; Hani, F.; Ali, H.M.; Ariffin, S.; Huda, N.; Zulkipli, N. Study of S-box Properties in Block Cipher. In Proceedings of the International Conference on Computer Communication and Control Technology, Langkawi Island, Kedah, Malaysia, 2–4 September 2014. [Google Scholar].

[13] Manjula, G.; Mohan, H.S. Constructing Key Dependent Dynamic S-Box for AES Block Cipher System. In Proceedings of the International Conference on Applied and Theoretical Computing and Communication Technology, Bengaluru, Karnataka, India, 21–23 July 2016. [Google Scholar].

[14] HafnerKarlheinz, C. Ritter Hartmut, M. Schwair Thomas, Wallstab Stefan, Deppermann Michael, GessnerIuergen, Koesters Stefan, Moeller Wolf-Dietrich, SandwegGerd, "Design and Test of an Integrated Crypto chip", IEEE Design & Test Of Computers, pp. 6-17, December 1991.

[15] R. Zimmermann, A. Curiger, H. Bonnenberg, H. Kaeslin, N. Felber, W. Fichtner, "A 177 Mb/s VLSI Implementation of the International Data Encryption Algorithm", IEEE JOURNAL OF SOLID-STATE CIRCUITS, vol. 29, no. 3, pp. 303-307, March 1994.

[16] X. Lai, J. L. Massey, "A proposal for a new block encryption standard" in Advances in Cryptology-EUROCRYPT’, Berlin, Germany:Springer-Verlag, vol. 90, pp. 389-404,1990.

[17] Sumeet Kaur, Savina Bansal & R. K. Bansal, “Steganography and classification of image steganography techniques”, IEEE Xplore: 12 June 2014.

[18] Chandramouli, R., Kharrazi, M. and Memon, N., “Image steganography and steganalysis: Concepts and Practice”, Proceedings of the 2 ndInternational Workshop on DigitalWatermarking, October 2003.

[19] Video Steganography by LSB Substitution Using Different Polynomial Equations ǁ, A. Swathi, Dr. S.A.K Jilani, Proc.International Journal of Computational Engineering Research (ijceronline.com) Vol. 2 Issue. 5.

[20] ShrutikaSuri, etal. “Comparative Analysis of Steganography for Coloured Images”, International Journal of Computer Sciences and Engineering, Volume-2, Issue-4.

[21] David Naccache, David MRaYhhi, "CRYPTOGRAPHIC SMART CARDS" in IEEE Micro, IEEE, pp. 14-24, June 1996.

[22] Gilles Brassard, Claude Crepeau, MiklosSantha, "Oblivious Transfers and Intersecting Codes", IEEE TRANSACTIONS ON INFORMATION THEORY, vol. 42, no. 6, pp. 1769-1780, November 1996. https://doi.org/10.1109/18.556673