Cryptographic approach to Securely Share and Protect Genomic Data

J. Santhosh Kumar[1], R. Vishnu[2], V. Ragulraja[3], M. K. Nivodhini[4], P. Vasuki[5]

[1, 2, 3] B. E., (CSE), Department of Computer Science and Engineering, KSRCE, TamilNadu, India.
Email: sashkumar11@gmail.com, viratvishnu1105@gmail.com, vragul0602@gmail.com

[4, 5] Assistant Professor, Department of Computer Science and Engineering, KSRCE, TamilNadu, India.
Email: nivodhinimk99@gmail.com, vasukiabi@gmail.com.

Abstract - At times it is significant to commune secret information to an individual or to a group of selected people and if it is intercepted and changed by an intruder may lead to undesired problems. To protect trusted information and to connect it to the person(s) concerned is a crucial task. One of the methods used for this is Cryptography that ciphers the evidence based on definite algorithm that makes it human unreadable unless decrypted in a predefined method set by the material sender. A large variation of cryptographic systems are used which have their own strengths and weaknesses. Digital data particularly image files are extensively used more internet. This paper is a try to give an outline of software data cryptography and cryptanalysis and employing disordered structure as possible tenacity for image encryption over customary cryptographic algorithms.

Keywords: Plaintext, Cipher text, Key, Encryption, Decryption

I. INTRODUCTION

Cryptography is a means for secret announcement where the messages are tousled through an encryption process to produce an unreadable cipher text that needs to undergo decryption to retrieve back the original message. It protects messages from unlicensed access where there is no access control. Various Cryptography terminologies are projected in Table-I. In this paper the section II confers the security requirements for cryptographic applications and various Cryptographic techniques based on key usage. Section III briefs about Pandemonium theory and its application to cryptography with special orientation to image encryption. Section IV is devoted on Cryptanalysis and Section V draws the termination and hope scope in this field.

II. CRYPTOGRAPHY- REQUIREMENTS AND TECHNIQUES

A. Requirements:

With respect to application based communication, there exists certain security requirements which includes: Validation (the method of ensuring the user's identity), Confidentiality (Ensuring that the message can be read by the intended user only), Consistency (Assuring the receiver that the message gets original message) and Non-refutation (A system to ensure that the sender really sent the message received by the user) [2, 3].

B. Types of Cryptography based on the Key Usage

It includes hash meanings, symmetric key cryptography and public key cryptography. The variation and there security analysis is discuss as below:

1. Hash Function

Also called meaning digest and one-way encryption, it uses a mathematical renovation to irreversibly encrypt information. Moderately than using keys; a fixed-length one-way hash charge is computed based upon the plaintext. It is well-suited for guaranteeing data integrity. Some of the common Hash algorithms include:

- Message Digest (MD) algorithms: It is an arrangement of byte-oriented systems that produce a 128-bit hash charge from an arbitrary-length message. The different version includes-
- MD2-It is defined in RFC 1319 and is planned for systems with small memory but, it is insecure against collisions attack (Rogier et al, 97) and preimage attack (Muller, 04).
- MD4: It is defined by RFC 1320 and planned exclusively for fast processing in software. The digest length is 128 bits. But its safety is breached by full collision attack (Dobbertin, 1995 and, Wang et al, 2004,) and theoretical preimage attack (Leurent, 2008).
- MD5: It is defined in RFC 1321. It is similar but slower to MD4 as more treatment is made to the original data. It produces a 128-bit hash value. Key usage includes checking data honesty. However, MD5 is not secure against collision attack.
- Secure Hash Algorithm (SHA): The three SHA algorithms are prepared in a different way and are SHA-0, SHA-1 and SHA-2.

- SHA-1: It is producing a 160-bit hash value and was originally published as FIPS 180-1 and is described in RFC 3174. Though, in 2005, security flaws were recognized in SHA-1.
- SHA-2: It describe four algorithms in the SHA: SHA-224, SHA-256, SHA-384, and SHA-512 which can construct 224, 256, 384, or 512 bits long hash values, respectively. SHA-224, -256, -384, and -512 are definite in RFC 4634. SHA-256 and SHA-512 are added with 32- and 64-bit words, in that order. It uses different shift amounts and additive constants but have similar structure reverse in number of iterations. SHA-224 and SHA-384 are abridged versions of the first two, figured with different initial values. The best public cryptanalysis shows attack breach pre-image resistance for 46 out of 80 rounds of SHA-512, and 41 out of 64 rounds of SHA-256.. Efforts are underway to develop enhanced alternatives and SHA-3, is currently under development.
- RIPEMD: It is a series of memo digests that initially came from the RIPE (RACE Integrity Primitives Evaluation) project. The 256 and 320-bit version reduces chance collision, but don't provide better security (against preimage attack) as compared to RIPEMD-128 and RIPEMD-160.
- HAVAL (Hash of Variable Length): Designed by Zheng et al, it is a hash algorithm with many planes of security. It can form hash values that are 128, 160, 192, 224, or 256 bits in length. HAVAL also agrees users to specify the number of rounds (3-5) to be used to cause the hash. The HAVAL hashes are signified by 32, 40, 48, 56 or 64-digit hexadecimal numbers. Though, HAVAL is no more secure after collision attack by Wang et al, 2004.
- Whirlpool: It operates on letters with length less than 2256 bits, and causes 512 bits hash. Though, in 2009 a recoil attack was broadcast that presents full impacts against 4.5 rounds of Whirlpool in 2120 operations making it doubtful.
- Tiger: Designed by Anderson et al, it is sheltered and works resourcefully on 64-bit workstations. Tiger-192 produces a 192-bit output. But, cryptanalysis attacks (Kelsey et al, Mendel et al, collision finding attack,) shows that it is no more protected.

2. Secret Key Cryptography (SKC)

It uses a one key for both encryption and decryption. The source uses the key to encrypt the plaintext and points the cipher text to the receiver. The handset applies the comparable key to decrypt the message and convalesce the plaintext. The main tricky with it is how to create, store and transmit key to those who need to decrypt communications. Mathematically, comparation pair (1-2) represents encryption and decryption process. It is obvious that the key must be known to both sender and receiver; that is the top-secret. It is considered into stream cryptographs and block encryptions. Stream ciphers operate on a single bit at a time and gadget some form of feedback mechanism so that the key is always changing. A block cipher outline encrypts one block of data at a time using the same key on each block. Here, the same plaintext block will always convert to the same cipher text when using same key in a block cipher. Block ciphers can operate in four styles viz. Electronic Codebook mode, Cipher Block Chaining mode, Cipher Feedback mode and Output Response mode. Some of the top-secret key algorithms are described as below:

Data Encryption Standard (DES): It is a block-cipher using a 56-bit key that works on 64-bit blocks. The number of rounds occupied is 16 and has a structure of Composed Feistel network. It employs Shift permute process for key generation and the mathematical operation is XOR. DES was designed exactly to yield fast hardware implementations and slow software employments. But, DES is insecure due to the 56-bit key size being too small and hence archaic. Further cryptanalysis shows that a brute energy attack is possible. As of 2008, linear cryptanalysis attack involves 243 known plaintexts (Junod, 2001).

Triple DES: It is a block cipher that smears the DEScipher system three times to each data block. The key size is 112/168 with 48 records of rounds and sub keys. The shift permute method is active for key generation and has a Feistel link structure. The cryptanalysis spells are 232 known plaintexts (Lucks) and 228 target keys with selected plaintexts per key (Biham). However, this is not presently practical and NIST studies it to be apposite.

Advanced Encryption Standard: It is based on asubstitution-permutation network but does not use a Feistel link[5]. AES has a fixed block size of 128 bits and a key scope of 128, 192, or 256 bits. AES has 10 rings for 128-bit keys, 12 rounds for 192-bit keys, and 14 rounds for 256-bit keys. It operates on a 4×4 column-major direction matrix of bytes. The AES cipher is computed as a number of repetitions of uprising rounds that convert the input plain text into the final output of ciphertext. The opening key-recovery attacks on full AES stayed by Bogdanov et al 2011. However, all known spells are computationally infeasible.

Blowfish is a keyed, symmetric block cipher. The keyproject features include key-dependent S-boxes and a highly intricate key schedule. Blowfish has a 64-bit slab size with variable key length ranging from 32-448 bits. It employs 16-round Feistel cipher and huge key-dependent S-boxes. The cryptanalysis shows that four rings of Blowfish are vulnerable to a 2nd -order differential attack (Rijmen, 1997).

3. **Public Key Cryptosystems (PKC)**

PKC be contingent upon the existence of one-way roles that are easy to compute whereas their inverse function is relatively difficult to compute. It has two different keys for data transmission. There exists a scientific relation between the two keys so that if one is used to encryption other can be used for decryption. It contains a Private Key (known only to owner) and a public key distributed to any user who requests it. Mathematically, the following equations represent the

encryption and decryption development where we have a key pair (k1, k2), k1 being public and k2 private. This way could be best used for non-repudiation. Some PKC are:

RSA: It employs Chinese Balance Theorem for key group while the mathematical procedure is factoring problem. RSA uses a variable size encryption block and a flexible size key (1024 to 4096). The cryptanalysis shows that a 768-bit key has been smashed. RSA is used in software goods, for digital signatures, key discussion, and encryption of small data slabs. Diffie-Hellman: It permits two parties to jointly start a shared secret key over an uncertain transportations channel that can be for encrypting ensuing communications using a symmetric key cipher. The Diffie–Hellman key promise provides the basis for valid protocols, and is used to provide perfect forward secrecy in Transport Layer Security's ephemeral methods.

Digital Signature Algorithm: The system specified in NIST's Digital Cross Standard, provides digital autograph capability for the certification of messages. Key group has two phases. The first time is choice of algorithm limits while the second phase subtracts public and isolated keys for a single user. Elliptic Curve Cryptography: A PKC process based upon elliptic curves over finite pitches. It offers levels of haven with small keys comparable to RSA. It is expected that finding the severed logarithm of a random elliptic curve element with respect to a publicly known base point is infeasible. Even though Triple-DES and IDEA etc. can achieve high safety, it is not appropriate for multimedia applications due to its large data sizes and real time constraint. In SKC, safe key conversation between two parties is a major constraint. Using PKC for encryption is very slow. Hence these systems are not up to mark and we need to find different resolution. A key to this can be to use confused encryption where, the encryption algorithm employs the pixels of an image instead of employing the bits of the image.

III.CHAOTIC CRYPTOGRAPHY FOR IMAGE ENCRYPTION

A. Theory of Chaos

It means "a state of disorder". It becomes non-linear if its limitation, internal variable, external signals, control variable, or even initial value is preferred in a specific way. This volatility of a deterministic system is termed as chaos. It is built on the fact that simple rules after iterated can give rise to complex behaviour. For a dynamical scheme to be chaotic, it must have the next properties:

1. It must be delicate to initial situations exponentially: It means that each point in such a system is illogically closely approximated by other points with different future trajectories.
2. It must be topologically mixing: It means that the scheme will change over time so that any given region or open set of its phase interplanetary will finally overlap with other region.
3. Its interrupted orbits must be dense: It means that every point in the space is advanced closely by interrupted orbits.

B. Similarities in chaotic maps and cryptography

This includes sensitivity to a modification in initial conditions and parameters, uneven periodic orbits with elongated periods and random-like behaviour. The distribution and confusion properties necessary in a cryptographic algorithm are attain through the iteration. The iterations of a chaotic map increase the initial region over the entire phase space .The parameters of the chaotic map may represent the key of the encryption algorithm. Chaotic systems are very sensitive to initial conditions and system limitations. For a given set of parameters in chaotic regime, two close initial conditions lead the system into deviating paths. Therefore encryption / decryption scheme can be found if the parameters are selected as “Keys” and “Trajectories” are used for the same. Since the same limitations are used for encryption and decryption, the chaos scheme is symmetric. The limitations and the initial conditions form a very large key space thereby enhancing parameters and the initial conditions form a very large key space thereby enhancing the security of the code. This review discusses some of the recent chaotic encryption techniques in brief.

C. Existing Methods for Chaotic Image Encryption

Baptista Method: Baptista uses logistic map in which the iterates are created using the equation. By taking the limit r for chaotic regime and with initial ailment x0 □ [0,1]. But, the security investigation shows following four defects in it. The circulation of the cipher text is non-uniform, the encryption is speed very deliberate, the cipher text size is more than the plaintext size and it is insecure against some different attacks. Hence, we need to look for some new system. Zhang et al (2011)proposed a system based on logistic map and cheat image somewhere he chooses the initial state and control stricture of logistic map as the top-secret key. But there exists softness, such as small key space making it insecure. Yong et al (2011) proposed extra image encryption method using PN Order in chaotic maps. Here a clandestine key is defined as initial settings for a chaotic map such as logistic map. The safety analysis shows that for any pixel of the plain image, encryption and decryption scheme is perverse and that decryption scheme is improper.

IV.CRYPTANALYSIS

It is the reverse process of cryptography. The objective of cryptanalyst is to be intelligent to decrypt cipher text.

A. Attacks on Key based Cryptography

- Cipher text Only Attack: Here the attacker obtains a sample of cipher text without the plaintext related with it.
- Known Plaintext Attack: The attacker obtains the model of cipher text and the corresponding plaintext.

- Chosen Plaintext Attack: The attacker can choose the quantity of plaintext and then get the corresponding encrypted cipher text.
- Adaptive Chosen plaintext attack: A cryptanalyst can mount this occurrence when he has decryption hardware but it is incompetent to abstract the decryption key from it.
- Brute Force Attack: Here key size delivers a lower bound on the security of the cryptosystem
- Related Key Attack: Here the invader can see the action of cipher under various keys whose values are originally unknown but where some mathematical bond involving the keys is known to the invader.
- Differential Attacks: This attack traces the modifications through changes discovering the cipher exhibiting non-random behaviour and manipulating them to recover secret key.

B. Hash Functions and Attacks

- Collision attack- It performances on a cryptographic hash by trying to find two random inputs that having same hash value.
- Pre-image attack is occurrence for finding a message that has a exact hash value.
- Birthday attack -The occurrence depends on the higher possibility of collisions found between random attack challenges and a fixed degree of permutations.
- Rainbow table is a pre-computed table for withdrawing hash functions, expressly for cracking key messes.
- Distinguishing attack –here the attacker can extract some info from encrypted data enough to distinguish it from random data.
- Side channel attack is based on material existing from physical operation of cryptosystem.
- Dictionary attack is a system for defeating a cipher by irritating to determine its decryption key or pass saying by searching likely options.

V.CONCLUSION

Cryptography is a powerful tool to protect information. In the recent centuries cryptography and cryptanalysis had seen a lot of exploration. Though due to varying supplies of applications and different types of digital data there does not exist a single cryptographic algorithm that could meet all supplies. Conventional cryptographic methods are appropriate for textual data however it is not suitable for Images and chaotic cryptography seems to be the best solution for image and video encryption since it is fast and computationally possible for large data sizes. However, computer are finite state machines and implementing true chaos on them is not likely. Implementing chaos for cryptography using logistic maps and difference equations are only solutions which do have their restrictions. Hence developing a fully protected chaotic encryption algorithm is still a challenge.

REFERENCES

[1] W. Stallings, Cryptography and link security: principles and practice: *Prentice Hall*, 2010.
[2] Faisal T, Ibrahim F, Taib MN (2010) A noninvasive intelligent approach for predicting the risk in Sleep Apnea sleep disorder. Expert Systems withApplication 37:2175-2181.
[3] J. Amigo, et al., "Theory and exercise of chaotic cryptography," Physics Letters A, 2007, vol. 366, pp.211-216.
[4] Zhang et al, "Implementation approaches for AES algorithm," Circuits and Systems Magazine, IEEE, 2003, vol. 2, pp. 24-46.
[5] "Chaos based cryptography: A newStyle to securetransportations",BARC,July-2005,http:// www.barc.gov.in/publications/nl/2005/200507-1.pdf
[6] Li et al, "Baptista-type chaotic cryptosystems: Worries and countermeasures, Physics Letters A, Elsevier Science, 2007, pp.368-375.
[7] Zhang Yong, "Image Encryption with Logistic Map and Cheat Image", Computer Research and Development (ICCRD), 3rd International Conference, IEEE, 2011, Vol. 1, pp. 97 - 101
[8] Gao et al, "A new chaotic algorithm for image encryption" Elsevier Ltd. 2005
[9] Yong Zhang "Comments on -An image encryption scheme with a pseudorandom permutation based on chaotic maps", Cross Strait Quad-Regional Radio Science and Wireless Technology Conference, IEEE Conferences, 2011 Vol. 2, pp. 1251 - 1255
[10] T. Xiang, et al., "A novel block cryptosystem based on reiterating a messy map," Physics Letters A, 2006, Vol. 349, pp. 109-115.
[11] Diffie, "The First Ten Centuries of Public-Key Cryptography" Proceedings of the IEEE, 1988, Vol.76 , pp. 560 – 577.