Digital Right Management Model based on Cryptography for Text Contents

Pushpendra Kumar Verma
Research Scholar, Department of CSE, Swami Vivekanand Subharti University, Meerut, U.P., India.

Dr. Jayant Shekhar
Professor, Swami Vivekanand Subharti University, Meerut, Uttar Pradesh, India.

Preety
Assistant Professor SIMC, Swami Vivekanand Subharti University, Meerut, U. P, India.

Abstract – Generally, DRM (Digital Rights Management) system is achieved with individual function modules of cryptography, watermarking and so on. In this typical system flow, all digital contents are temporarily disclosed with perfect condition via decryption process. This paper describes the fundamental idea of a novel DRM method which is composed of an incomplete cryptography and user identification mechanism to control the quality of digital contents. This research paper proposes an improved scheme to encrypt the plain text message for its security. All the conventional encryption techniques are very weak and brute force attack and traditional cryptanalysis can be used to easily determine the plain text from encrypted text. In this work of encryption technique, a new concept of conventional ceaser cipher algorithm with hill cipher algorithm is used to make encryption technique more secure and stronger than the earlier concept. The plain text is encrypted in such a way that it becomes difficult to decrypt it. The proposed system is divided into two phases. In first phase, the plain text message is converted to first encrypted text using a new substitution approach which uses poly-alphabetic cipher technique. The encryption is done using variable length key which depends on the string length. In the second phase, the hill cipher technique is applied on the first encrypted text to produce new encrypted text or cipher text. At the receiver end, if the receiver has appropriate decryption key, he can generate the text message similar to the original message. DRM is a form of cryptography, the process of protecting information from unauthorized use by transforming it so that only the authorized receiver can read it. The sender – in this context, the information vendor or provider – encrypts the digital object via a "key" of some kind. The recipient – the information user – may decrypt it for use with a copy of the same key, usually automatically. The information is protected from "attackers" – unauthorized users, or uses – who lack the key.

Encryption Standard (AES). Features of data are depends on its types. Therefore same encryption technique cannot be used used for all types of data. Images have large data size and also has real time constrain problem hence similar method cannot be used to protect images as well as text from unauthorized access. However with few variations in method AES can be used to protect image as well as text. In this project I have implemented encryption and decryption for text and image using AES.

Index Terms – Key Based Interval Splitting (KSAC), Advanced Encryption Standard (AES).

1. INTRODUCTION

Chosen plaintext for Randomized Arithmetic Code are based on same key used to encode many messages, known that using a same key for many messages leads to insecure encryption scheme .The strongest version of security is chosen-cipher text security where the adversary has access to the encryption and decryption engine (but does not know the secret key) and can decrypt any cipher text of his choice or encrypt any plaintext to his choice. Every message will be compressed using a new key sequence achieved using a secure pseudorandom sequence generator. Arithmetic coding followed by XOR with a secure Pseudorandom bit sequence leads to an encryption scheme that is chosen plaintext secure. The XOR can be incorporated into AC thereby incurring minimal penalty for real-time applications. RAC that uses different key for different messages is not secure under cipher text only attack, but secure in chosen plaintext attack. When both compression and security are sought, one approach is to simply use a traditional arithmetic coder in combination with a well-known encryption method such as the Advanced Encryption Standard (AES)

1.1 Terminologies in Cryptography

An encryption technique has five ingredients [2]:

1. Plaintext: This is the original intelligible message or data that is fed into the algorithm as input.

2. Encryption Algorithm: The encryption algorithm performs various substitutions and transformations on the plaintext.

3. Key: The key is also input to encryption algorithm. The key is a value independent of the plaintext and of the algorithm. The algorithm will produce a different output depending on the specific key being used at the time.

4. Ciphertext: This is the scrambled message produced as output. It depends on the plaintext and secret key.

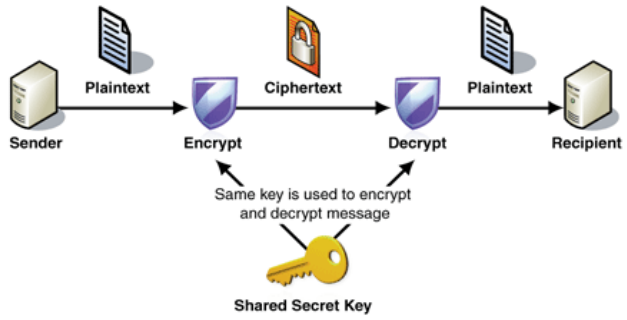
5. Decryption Algorithm: This is essentially the encryption algorithm run in reverse. It takes the ciphertext and the secret key and produces the original plaintext. [10][11]

1.2 Types of Cryptography

Cryptography is a technique in which secret messages are transferred in the encrypted form from sender to receiver over the communication line. Cryptographic techniques are very useful to protect secret information. They protect the secret or confidential information by converting the information to some unintelligible form using a key. To retrieve the information, the encrypted information should be converted back to original information using some keys. Based on the key, the cryptography can be classified into two categories [1]:

1. Shared key cryptography
2. Public key cryptography

Shared key cryptography also called symmetric key cryptography or private key cryptography or secret key cryptography in which, same key is used for encryption and decryption i.e. both the sender and the receiver know the same key. Ex. DES, 3 DES, AES, etc. Messages are encrypted by the sender using the key and decrypted by the receiver using the same key.



Encryption Decryption Fig. 1: Shared Key Cryptography

Figure 3 shows process of secret key cryptography [2]. Here same key is shared by both sender and receiver for encryption and decryption. Public key cryptography also called asymmetric key cryptography which uses different keys for encryption and decryption. Ex. RSA, Digital signature scheme, etc. The public key is known to all the receivers, is used for encrypting the plaintext message. The private key is known only to the user of that key. With public key cryptography, keys work in pairs of matched public and private keys. Figure 4 shows the process of public key cryptography where public key used by sender for encryption and all the receivers use their private keys for decryption. Messages encrypted using the public key cannot be decrypted using the public key. Public key encrypted messages can only be decrypted using corresponding private key which is kept secure. Asymmetric key cryptography is very slower and has very higher computational costs which are most of the time prohibitive for multimedia data. Symmetric key cryptography is fast, comparatively lower cost and may be used for multimedia data. Encryption Decryption

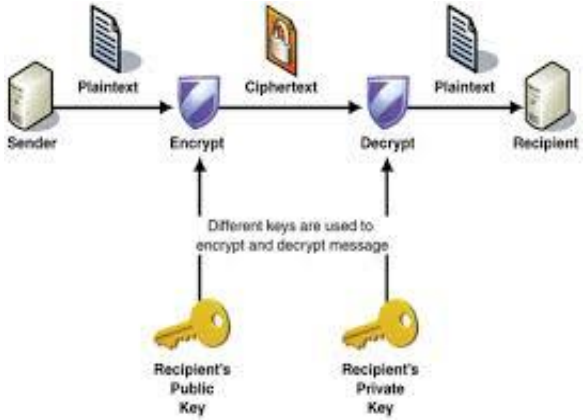


Fig. 2: Public Key Cryptography Hash function is another type of cryptography which makes use of some mathematical transformation. [10]

1.3. Hill Cipher

The Hill cipher (HC) algorithm is one of the famous and known symmetric algorithms in the field of cryptography. It is a poly-alphabetic cipher proposed by the mathematician Lester Hill in 1929 in the journal of mathematics. Hill cipher requires a matrix based polygraphic system [5]. The Hill cipher takes m successive plaintext characters and substitutes for them m ciphertext characters. The core of Hill cipher is matrix manipulations. For encryption, algorithm takes m successive plaintext characters and instead of that substitutes m cipher characters. For example m=2; {a b c d e f…}=ab cd ef… or abc def… and so on. In Hill cipher, each character is assigned a numerical value like a = 0, b = 1, ... , z = 25. The substitution of ciphertext characters in the place of plaintext characters leads to m linear equation. For m = 3, the system can be described as follows: C=KP Plain Text Cipher Text Cipher Text Plain Text where C and P are column vectors of length 3, representing the plaintext and ciphertext, respectively and K is a 3 x 3 matrix, which acts as key for encryption. All operations performed with modulus of 26. In Hill cipher key is an invertible m x m matrix, where m is block length. Decryption process uses inverse of matrix K. The inverse matrix K-1 of a matrix K is defined by following equation. KK-1 = K-1K=I where, I is the Identity matrix. But the inverse of matrix does not always exist and when it exists, it satisfies above equation. The inverse matrix K-1 is used to decrypt the ciphertext. In general it can be written as follows: Encryption Process: C=Ek (P)=Kp Decryption Process: P=Dk (C)=K-1C= K-1Kp=P If the block length considered as m, there are 26m different m characters blocks are possible.

2. LITERATURE SURVEY

This section consists of brief description of few classical and modern plain text encryption techniques.

2.1 Caesar Cipher:

The Caesar cipher is one of the earliest known and simplest ciphers. The method is named after Julius Caesar, who apparently used it to communicate with his generals. It is a type of substitution cipher in which each letter in the plaintext is 'shifted' a certain number of places down the alphabet. For example, with a shift of 1, A would be replaced by B, B would become C, and so on. The Caesar cipher offers essentially no communication security, and it can be easily broken even by hand.

2.2 Playfair Cipher :

The Playfair cipher is a polygraphic cipher which enciphers two letters at a time. The Playfair cipher was the first practical digraph substitution cipher. The scheme was invented in 1854 by Charles Wheatstone, but was named after Lord Playfair who promoted the use of the cipher [4]. The technique encrypts pairs of letters, instead of single letters as in the simple substitution cipher. The Playfair is significantly harder to break since the frequency analysis used for simple substitution ciphers does not work with it. Frequency analysis can still be undertaken, but on the 25*25=625 possible digraphs rather than the 25 possible monographs. Frequency analysis thus requires much more ciphertext in order to work. It was used for tactical purposes by British forces in the Second Boer War and in World War I and for the same purpose by the Australians during World War II. This was because Playfair is reasonably fast to use and requires no special equipment. A typical scenario for Playfair use would be to protect important but non-critical secrets during actual combat. By the time the enemy cryptanalysts could break the message the information was useless to them.

2.3 Affine Cipher

The Affine cipher is a special case of the more general monoalphabetic substitution cipher. The cipher is less secure than a substitution cipher as it is vulnerable to all of the attacks that work against substitution ciphers, in addition to other attacks. The cipher's primary weakness comes from the fact that if the cryptanalyst can discover (by means of frequency analysis, brute force, guessing or otherwise) the plaintext of two ciphertext characters, then the key can be obtained by solving a simultaneous equation.

2.4 Vigenere Cipher

Vigenere cipher was proposed by Blaise de Vigenere in the 16th century. Vigenere cipher is poly-alphabetic substitution cipher in which a single plain text letter can be converted into multiple cipher text letters. This conversion depends on the position of the letter in the plain text e.g. c may be converted into g because it is at position 3 in the plain text but c can be changed into z because its position in the plain text is 5. Vigenere cipher makes use of Vigenere table of size 26 X 26.

2.5 Blow-Fish Cipher

Blowfish cipher was developed by Bruce Schneier. Blowfish is a variable-length key, 64-bit block cipher. The algorithm consists of two parts: a key-expansion part and a dataencryption part. Key expansion converts a variable-length key of at most 56 bytes into several sub key arrays totaling 4168 bytes. Data encryption occurs via a 16-round Feistel network. Each round consists of a key-dependent permutation, and a key- and data-dependent substitution.

2.6 Rail Fencing Cipher

Rail fencing technique involves: Writing plain text message as a sequence of diagonal and reading it as a sequence of row to produce cipher text. In a Rail Fence cipher, after removing the spaces from the original message, write the characters in the message in the zig-zag pattern. The key for the Rail Fence cipher is just the number of rails.

2.7 Modern Ciphers

Modern ciphers use both substitution and transposition to encrypt the message that increases the security of data. The data is encrypted in blocks instead of single characters at time. The well-known example of block cipher is Data Encryption Standard (DES). DES uses 56-bits key and encrypt 64-bits of data as a single block. AES uses the same key for both encryption and decryption. The AES standard states that the algorithm can only accept a block size of 128 bits and a choice of three keys - 128, 192, 256 bits. Depending on which version is used, the name of the standard is modified to AES-128, AES-192 or AES- 256 respectively.

The Rijndael Algorithm is the new Advanced Encryption Standard (AES) recommended by the US National Institute of Standards and Technology (NIST) for protecting sensitive, unclassified government information. Since Rijndael is an iterated block cipher, the encryption or decryption of a block of data is accomplished by the iteration of a specific transformation. As input, Rijndael accepts one-dimensional 8-bit byte arrays that create data blocks. The plaintext is input and then mapped onto state bytes. The cipher key is also a one-dimensional 8-bit array .With an iterated block cipher; the different transformations operate in sequence on intermediate cipher results.

The columnar transposition cipher is a fairly simple, easy to implement cipher. It is a transposition cipher that follows a simple rule for mixing up the characters in the plaintext to form the ciphertext. Although weak on its own, it can be combined with other ciphers, such as a substitution cipher, the

combination of which can be more difficult to break than either cipher on its own.

2.8 Cryptanalysis

In Cryptanalysis, the attacker uses various methods to get the plain text from the cipher text. They try to find out the way in which plain text is converted into cipher text and the encryption key used. Various methods were used for identifying ciphers. Identification of permutation, substitution and Vigenère ciphers was done using frequency analysis. An attempt was made to identify block ciphers like DES and Blowfish using pattern recognition method. Other ciphers like stream cipher SEAL and Enhanced RC6 have been identified using neural networks.

3. OBJECTIVES

The core objective of this paper is to protect information leakage what so ever manner it may be, the use of appropriate technology. To provide a high level of confidentiality, integrity, non-reputability and authenticity to information that is exchanges over networks.

Confidentiality – data is protected by hiding information using encryption technique.

Integrity – Ensures that a message remains unchanged from the time it is created and opened by recipient.

Non – reputability – it provide a way of proving that the message came from someone even if they try to deny it. Authentication – it verifies the identity of user in the system and continues to verify their identity in case someone tries to break into the system.

4. PROPOSED SCHEME

The proposed scheme, improve the security of Key Based Interval Splitting (KSAC) using Advanced Encryption Standard (AES).

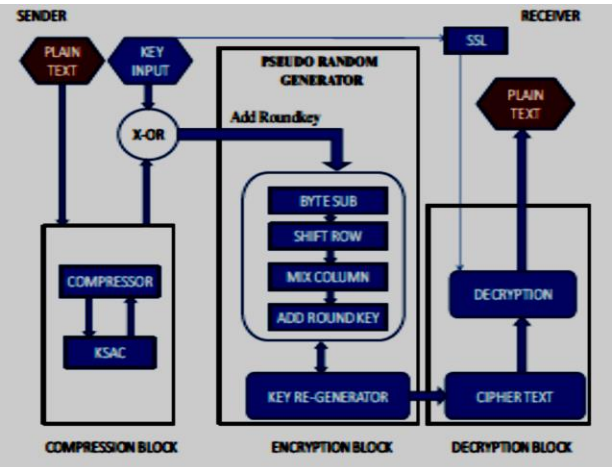


Fig.3. Proposed System Architecture

A. Compression Block

The plain text is the input to the compressor. This plain text gets compressing by key based interval splitting arithmetic coding. The actual compression is performed by the interval splitting arithmetic coder. In an interval splitting AC, the intervals associated with each symbol, which are continuous in a traditional arithmetic coder, can be split according to a key known both to the encoder and decode data compression, source coding, or bit-rate reduction involves encoding information using fewer bits than the original representation. Compression can be either lossy or lossless. Lossless compression reduces bits by identifying and eliminating statistical redundancy. No information is lost in lossless compression. Lossy compression reduces bits by identifying marginally important information and removing it. The process of reducing the size of a data file is popularly referred to as data compression, although it‟s formal name is source coding. Compression is the useful because it helps reduce resources usage, such as data storage space or transmissions capacity. Because the compressed data must be decompressed to use, this extra processing impose computational or other costs through decompression, the situation is far from being a free lunch. Data compression is the subject to a space-time complexity trade-off. For instances, a compression scheme for video may require expensive hardware for the video to be decompressed fast enough to be viewed as it is being decompressed, and the option to decompress the video in the full before watching it may be inconvenient or require additional storage. The design of data compression schemes involve trade-offs among the various factors, including the degree of compression, the amount of distortion introduced the computational resources required to compress and uncompress the data.

B. Encryption Block

In cryptography, encryption is the process of encoding messages in such a way that eavesdroppers or hackers cannot read it, but that authorized parties can. In an encryption schemes, the message or information is encrypted using an encryption algorithm, turning it into an unreadable ciphertext. This is usually done with the use of an encryption key, which specifies how the message is to be encoded. Any adversary that can see the ciphertext should not be able to determine anything about the original message. An authorized party, however, is able to decode the ciphertext using a decryption algorithm that usually requires a secret decryption key, that adversary does not have the access to. For technical reasons, an encryption scheme usually needs a key-generation algorithm to randomly produce keys. There are two basic types of the encryption schemes: Symmetric-key and public-key encryption. In symmetric key schemes, the encryption and decryption keys are the same. Thus communicating parties must agree on a secret key before they wish to communicate. In public-key schemes, the encryption key is published for anyone to use and

encrypt messages. However, only the receiving party has access to the decryption key and is capable of reading the encrypted messages. Public-key encryption is a relatively recent invention: historically, all encryption schemes have been symmetric-key schemes. The input key sequence is given to pseudorandom generator. The pseudorandom generator operations are bytesub, shift row, mixcolumn and addroundkey. Each of the bytes in the matrix changed to another byte in byte substitution. The rows of the matrix are shifted cylindrically to the left. The output of the shift row is multiplied by binary matrix. The add round key is derived by XOR with mixcolumn. Key regenerator gives the key for each key based interval splitting in arithmetic coding. Then it gives as input to the pseudo random generator until the given input splitting finished.

C. Decryption Block

When receiver receives the encrypted message, called cipher text he changes it back to plain text using a decryption key. Chosen plaintext means hacker gainstemporary access to the encryption machine. Receiver can encrypt a large number of suitably chosen plaintext and try to use the resulting cipher text to deduce the key. Chosen cipher text strings of symbols and try to use the results to deduce the key. No information is lost in lossless compression. Lossy compression reduces the bits by identifying marginally important information and removing it. Finally the original file which was send by the sender is decrypted and decompressed. Each type of data-compression algorithm minimizes redundant data in a unique manner. For example, the Huffman encoding algorithm assigns the code to characters in a file based on how frequently those characters occur.

5. CONCLUSION

In existing system Randomized Arithmetic coding is used to encrypt the message. In our proposed system we used Key based interval Splitting by the algorithm Advance Encryption Standard. It encrypts by the 128 bit key. For encryption we can use 256 and other combination of keys which is compatible for Advance encryption standard. And also we can change AES algorithm for encryption in future. The decryption method depends on the encryption method used.

REFERENCES

[1] Gary C. Kessler, “An Overview of Cryptography”, 2013.
[2] Ramandeep Sharma, Richa Sharma, Harmanjit Singh, “Classical Encryption Techniques”, International Journal of Computer & Technology, Vol. 3, No.1,August 2012, pp.84– 90.
[3] “CRYPTOGRAPHY”,https://en//.wikipedia.org/wiki/cryptography
[4] Henk C. A. van Tilborg, “FUNDAMENTALS OF CRYPTOLOGY”, pp. 9 - 21.
[5] M. Nordin A. Rahman, A.F.A. Abidin, Mohd Kamir Yusof, N.S.M. Usop, “Cryptography: A New Approach of Classical Hill Cipher”, IJSA, Vol. 7, No. 2, March 2013.
[6] Kashish Goyal, Supriya Kinger “Modified Caesar Cipher for Better Security Enhancement”, International Journal of Computer Applications, Vol. 73, No. 3, July 2013.
[7] Maneesh Upmanyu, Anoop M. Namboodiri, Kannan Srinathan, and C. V. Jawahar, “Blind Authentication: A Secure Crypto-Biometric Verification Protocol”, IEEE Transaction on Biometric, Vol. 5, No. 2, June. 2009.
[8] A. K. Jain, A. Ross, and S. Prabhakar, “An introduction to biometric recognition,” IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 1,pp. 4–20, Jan. 2004.
[9] Katz, J. and Lindell, Y. (2008) „Introduction to Modern Cryptography‟, London, U.K.: Chapman & Hall/CRC.
[10] Pushpendra Kr. Verma , Dr. J. Shekhar ,Preety and Amit Asthana“ Digital right management to foster mobile multimedia services” in IJMIE ISSN No. 2249-0558, Vol. 2 Issue 4,pg 246-353, April 2012. http://www.ijmra.us.
[11] Pushpendra Kr. Verma , Dr. J. Shekhar ,Preety and Amit Asthana“ Digital right management to foster mobile multimedia services” in ZENITH International Journal of Multidisciplinary Research ,ISSN 2231-5780 Vol.4 (1), pg 9-20 April 2012, JANUARY (2014). www.zenithresearch.org.in .
[12] C. Shi, S.-Y.Wang, and B. Bhargava, “MPEG video encryption in realtime using secret key cryptography,” in Proc. 1999 Int. Conf. Parallel and Distributed Processing Techniques and Applications (PDPTA‟99), Las Vegas, NV, Jun./Jul. 28–1, 1999.
[13] H. Cheng and X. Li, “Partial encryption of compressed images and video,” IEEE Trans. On Signal Process., vol. 48, no. 8, pp. 2439–2451, Aug. 2000.
[14] M. Van Droogenbroeck and R. Benedett, “Techniques for a selective encryption of uncompressed and compressed images,” in Proc. Advanced Concepts for Intelligent Vision Systems (ACIVS) 2002, Ghent, Belgium, Sep. 9–11, 2002.

Author

Pushpendra Kumar Verma is pursuing in Ph.D.(CSE) from Swami Viveknand Subharti University Meerut.