World of Computer Science and Information Technology Journal (WCSIT)
ISSN: 2221-0741
Vol. 1, No. 4, 119-123, 2011

A Novel Approach for Security Development for Multimedia System

[1]B.Prabhavathi, [2]P.D.Chidhambara Rao, [3]P.Raja Sekhar, [4] Ishrath Jahan, [5] G.Murali Mohan
[1]Asst. Professor, AVR&SVR College of Engg.& Tech., Dept .of CSE, Nandyal, Kurnool(Dt.), A.P.
[2]Asst. Professor, SKTRM College of Engg. & Technology, Dept .of CSE, Kondair, A.P.
[3]Asso. Professor, SVPCET, Dept .of CSE, R.V.S.Nagar, Puttur, Chittur(Dt.), A.P.
[4]Asst. Professor, College for Charted Accounts, Dept. of CSE , Vijayawada, A.P.
[5]Asso. Professor, SVPCET, Dept .of CSE, R.V.S.Nagar, Puttur, Chittur(Dt.), A.P.

Abstract— Efficient multimedia encryption algorithms play a key role in multimedia security protection. We introducing multiple Huffman Tables (MHT), which performs both compression and encryption by using multiple statistical models (i.e. Huffman coding tables) in the entropy encoder and multiple Huffman tables are kept secret. A known-plaintext attack is presented to show that the MHTs used for encryption should be carefully selected to avoid the weak keys problem.

We then propose chosen-plaintext attacks on the basic MHT algorithm as well as the advanced scheme with random bit insertion. In addition, we suggest two empirical criteria for Huffman table selection, based on which we can simplify the stream cipher integrated scheme, while ensuring a high level of security.

Keywords- Cryptanalysis; Encryption; Entropy encoding; Multiple Huffman Tables (MHT); Selective Encryption.

I. Introduction

Digital rights management is a rapidly emerging area of research that deals with all aspects of secure data communication, from the system level key exchange protocols to the signal processing and ciphering algorithms employed to make the contents unusable by unauthorized parties. In particular digital rights management requires multimedia securization technologies, and enables applications such as copyright protection, authentication, and conditional access, just to mention few. Encryption is the one of the major digital rights management enabling technologies. Usually to provide confidentiality, the data is encrypted using a stream cipher or block cipher (Ex: DES[l] (or) AES [2]) in some mode of operation for encryption [3].Data encryption can be used to cipher all or parts of content. So that only the user that has received a key can decrypt and display ail or parts of the data. However, unlike the ordinary computer application, multimedia applications generate large amount of data that has to be processed in real- time. Hence, a number of techniques for realtime encryption of multimedia data have been proposed in past years. Two common approaches to real time multimedia data encryption are

1 .Selective encryption

2. Entropy coding that provides encryption.

The design philosophy of selective encryption is to provide faster encryption by encrypting only small portions of the multimedia data. Without the knowledge of the encrypted data, the adversary will not be able to recover the original data(e.g., image or video).since traditional schemes can be used for encryption. The issues related to selective encryption are more signal processing than cryptography related. The major issue is to select important information that will be encrypted. i.e the information whose encryption will generate that the adversary can't recover useful information about original data.

In the second approach, entropy coding provides encryption; the entropy coder has two functionalities: compression and encryption. The goal is to improve the efficiency by doing both compression and encryption in a single step. The entropy coders that provide encryption use secret keys to encode data. The adversary should not able to decode the data without the secret key. This approach can be combined with selective encryption for greater efficiency. However the security of most of the selective encryption systems is not high and often coding efficiency is sacrificed. Consequently designing a good multimedia encryption scheme that features a high level of security and low computational cost is challenging task.

A promising direction of research in this field, pioneered by Wu and d Kuo, is to combine encryption with entropy coding by using multiple statistical models [4] in the entropy codec. Entropy coding is the last stage of many multimedia compression systems, where a symbol stream is converted into a bit-stream via a statistical model. The major advantage of this scheme is that encryption is carried out at the same time of entropy coding, thus demanding only a negligible amount of computation. The high semantic security is guaranteed in the sense that, without knowing the key, the bit-stream cannot be correctly decoded. Even though the multiple Huffman tables (MHT) method proposed by Wu and Kuo has many desirable properties, it is vulnerable to the chosen plaintext attack (CPA). An Advanced MHT encryption scheme is proposed in this work to overcome this drawback.

To improve the security, several kinds of advanced MHT schemes have been proposed by either inserting random bit in the encrypted bit stream or integrating with a stream cipher [4], [5]. Recently, another scheme via random rotation in partitioned bit streams has been reported [6]and has been applied to a MHT system [7].

In this paper, we analyze the security of the basic MHT scheme and proposed advanced method. We show that the Huffman tables should be carefully selected; otherwise, the computations of exhaustive search will be significantly reduced. We also present chosen-plaintext attacks on the basic MHT and the advanced method with random bit insertion. Further, we suggest two empirical criteria for Huffman table selection, based on which we can simplify the stream cipher integrated scheme, while ensuring a high level of security.

The rest of this paper is organized as follows. Section II presents an overview of MHT methods. In Section III, we show the security analysis. Two empirical criteria for Huffman table selection, together with a simplified stream cipher integrated scheme, are suggested in Section IV, Section V is the conclusion

II. MHT ENCRYPTION SCHEMES

Encryption and compression are, in fact, intimately related in that it is redundancy in the data permits encryption and compression [8]. This fact motivates us to integrate encryption with compression. Huffman coder is very popular in modern multimedia compression system, with the aim of compressing symbols such as quantized DCT coefficients into bit streams, according to some predefined statistical models. Since, usually, the statistical model space is not large; it does not offer enough security. The MHT algorithm, in order to increase the model space while maintaining the computational efficiency, keeps the structure of the Huffman tree but enlarges the model space through mutating the original trees.

The procedure of basic MHT algorithm can be described as follows:

a) Train four original Huffman trees from different sets of training data. An example for JPEG dc coefficient coding can be found in Fig. 8 in [1],

b) Perform tree mutation to create the whole Huffman tree space. The operation is illustrated in Fig. 1.

c) Randomly select m different trees from the space, and number them from 0 to m-1.

d) Generate a random vector P={ p0…...pn_1 },where each p is an integer ranging from 0 to m - 1.)

e) For the i th symbol, use tree I mod n p to encode it

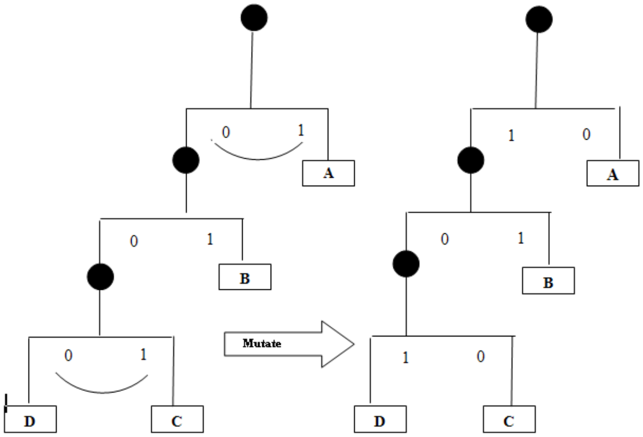

Figure 1. Tree Mutation Process. A, B, C, D denote alphabets.

Typically, we set m=8 and n=128. Since both the creation of Huffman tree space and the lookup table operations are quite simple, the MHT scheme needs small key-setup cost and encryption/decryption cost. On the other hand, due to the fact that decoding a Huffman coded bit stream without knowing the probability mass function of the source is surprisingly difficult[9], the basic MHT is claimed to be secure under cipher only attack and known-plaintext attack [4], [5]. However, under chosen-plaintext attack, it is vulnerable, even when the cipher can receive symbols as a whole chunk and output the corresponding code words all together [4]. Aiming to increase the security against chosen-plaintext attack, two types of enhanced schemes, selective random bit insertion scheme and stream cipher integrated scheme, have been proposed [4], [5]. In the former case, another random vector is generated, where is a one-bit integer [4]. For the i th bit in the encrypted bit stream, where is a constant, function will be performed do nothing for add random bit after for (l) For the stream cipher integrated scheme, a key stream is generated using a stream cipher, where is the seed [6]. Instead of XORing with the plaintext, the key stream will be partitioned into blocks, where is a -bit integer determining a Huffman tree to encode the encountered symbols.

We proposed encryption scheme via random rotation in partitioned bit streams was proposed by altering the generated bit stream [6]. Suppose X= xl,x2, x3 xN is a bit stream of N bits. This algorithm consists of the following two major steps, where pi and ri serve as the secret keys.

Partition X into k blocks Xi with length pi, i=1… k.

Perform a ri -bit left rotation on each block i A.

III. SECURITY ANALYSIS

A. KNOWN-PLAINTEXT ATTACK AND WEAK KEYS PROBLEM :

In this subsection, we propose a detailed analysis of known plaintext attack on basic MHT, and show its weak keys problem. Suppose we are given a bit stream X=Xl,X2,X3…….XN and know that it is the multiple Huffman encoding of some alphabet string C=cl,c2, …… cs . Since any bit stream is of equal probability, we have no knowledge to determine the boundary of X, i.e. establish the alphabet-codeword relationship. Hence, in order to partition X into S nonempty sub-blocks d1, d2, ds, we have to consider C sN-1 cases. In the traditional case where just one single Huffman table is used, we can check the prefix condition and consistence condition to eliminate some impossible partitions. In the current MHT scheme, we cannot do so since these two conditions are not necessarily satisfied for the whole bit stream. However, if we can estimate (he upper bound K to the length of a codeword, the number of partition can be reduced from C sN-1 =O (NS) to O(Ks). According to the Huffman tree creation procedure in Fig. 1 .we know K=11. Nonetheless, an attack with complexity of O (lls) is still formidable.

A more sophisticated method is to exploit the knowledge of the tree structure, i.e., the possible lengths of symbol. In the extreme case where all m the selected trees come from the same original tree, the identical symbol in the symbol string has the same length. From Table 1, we notice that any symbol takes at most three different lengths. This observation immediately reduces the number of cases to be checked to 3| Ω | =313, since we only need to guess the length once for all identical symbols. Here, Q denotes the symbol set {'O',’1’, .11','error'}, and | Ω | is the set size. The complexity of order 3|Ω| is already feasible, especially when the given plaintext does not include all | Ω | different letters.

TABLE I. POSSIBLE CODE LENGTH FOR SYMBOLS. IN “LENGTH” COLUMN FOUR NUMBERS CORRESPOND TO FOUR ORIGINAL TREES

Symbol	Length	Symbol	Length
‘0’	2 2 2 1	‘7’	5 7 6 6
‘1’	3 2 2 3	‘8’	6 8 7 7
‘2’	3 2 3 4	‘9’	7 9 8 8
‘3’	3 3 3 4	‘10’	8 10 9 9
‘4’	3 4 3 3	‘11’	9 11 10 10
‘5’	3 5 4 4	‘error’	9 11 10 10
‘6’	4 6 5 5	-----	------

In both basic and enhanced MHT, the Huffman trees are randomly selected from the tree space. The probability that all selected trees are mutated from the same original tree is:

P1=4 ∏ i=0 m-1 (M-i)/(4M-i) ~ 6 * 10-5 ……..(2)

Where M=212 is the total number of trees generated from one original tree. It is almost four times higher than that of the optimal case where all m selected trees are uniformly mutated from the four original trees

P2=M4(M-l)4 / 4 ∏ m-1 (4M-1)~1.52 * 10-5………..(3)

Note, also, that if all the m trees come from the last two original trees, around 2/3 of all alphabets in Ω will have fixed length. The occurring probability is:

P3= ∏ i=0 m-1 (2M-i) / 4M-i) ~ 0.0039 .... (4)

Hence, random selection of Huffman trees will introduce weak keys problem in the sense that some of the selected trees provide lower security than the others. To avoid this problem, a better solution is to randomly select ml 4 trees among each original tree space of size M , instead of choosing m trees from the whole space of size 4M.

B. CRYPTANALYSIS AGAINST THE SCHEME WITH RANDOM BIT INSERTION

Before presenting the cryptanalysis on the enhanced scheme inserting random bits, we propose an efficient chosen-plaintext attack on the basic MHT method.

Recall in [5], the author proved, under chosen-plaintext attack, the basic MHT could be broken in n | Ω | times of encryption oracle access. But this conclusion is only valid when the cipher receives one single symbol and outputs its corresponding codeword. In this work, we consider the more complicated case where a whole chunk of alphabets are received and the corresponding code words are output all together. We have the following proposition.

Proposition: Under chosen-plaintext attack, the basic MHT can be broken in n2 | Ω | oracle accesses, when the cipher outputs the bit stream of n alphabets all together.

Proof: The attack method can be described as follows.

Step 1: input alphabet stream c1,c2,c3……Cn-1 a0 and obtain bit stream Z1,Z2,….Zn-l, Zn Where Belongs to Ω can be any alphabet, ai is the ith alphabet and Zi is the corresponding codeword.

Step 2: Input another length -n alphabet stream cl,c2,c3,…cn-1, all and the output sequence is Z1,Z2,….Zn-l, Zn we choose all is because, from Fig. 8 in [1], a0 and all are separated by the root node, i.e., the first bit of the codeword co representing a0 and all is different. Therefore, we can obtain

Z’=Z1,Z2 Zn-l,Zn OO....O Ex-ORZl,Z2, Zn-l,Z'n
Denote the index of the first ‘1' Z as j (left to right order). Obviously, the codeword of a0 and all starts from the jth bit.

Step 3: Input length-n alphabet stream cl,c2,c3, ..cn-1.,a1. Since the start position has been determined in Step 2, it is straightforward to find the corresponding codeword of a1.

Continuing in this fashion, in Step |Ω| . all the Alphabet codeword pairs of the nth tree are recovered. Repeat such steps, we can restore all the alphabet-codeword pairs for all n trees, and the computational cost is n*n |Ω| =n 2|Ω| .

Now the key thing for the random bits insertion scheme is to find the locations of the inserted bits. Note that these locations have been fixed as long as the vector Q is determined. Assume the encoded bit stream of cl,c2,….cn is B=bl, b2,….bf yl bf+1…… bgy2 bg+1….bhy3bh+1,where bi is the original bit, and yi is the inserted random bit taking 0 or 1 with the same probability .Therefore, we can encode c1,c2, … cn T times, and obtain Bl, B2,… BT. We then perform the operations Di= B1 Ex-OR Bi, i = 2 ~ T. Due to the randomness of yi, the bit 1 in Di reveals where the inserted bits locate. The missing detection probability for every random bit is 2 -T. After successful detection, we simply remove them from the cipher, and apply the attack described in the proof to break the enhanced scheme.

Since when detecting the random bits, we exploit the random property, one may argue that what if inserting deterministic bits instead of random bits. In fact, this does not help much to improve the security. Recall that the random bits will be added after the (w*i) th bit if qI mod v =1 . Hence, in the bit stream available, the (w+1)th bit is a suspicious bit to be the first added bit, and (2w+l) th, (2w+ 2) th bit are two suspicious bits to be the second added bit. Along this line, we can find all suspicious bits. And according to [1], the number of random bits should not exceed 1% of the original cipher text, thus, the total number of suspicious bits is quite limited. For example, for a cipher with length 700, and w =70 , the total number of suspicious bits is 45. Consequently, after recovering all the alphabet codeword pairs for each tree, using the above mentioned method, we further perform two additional checking steps:

Step a: Check whether the obtained code words contain suspicious bits according to the suspicious bits index. If yes, go to Step b, otherwise, continue to find the alphabet codeword pairs for the previous tables.

Step b: Check whether the following equality holds, where \i denotes the length of obtained codeword of ai.

E= Σ i=0 |Ω| -1 2 –li = 1 ……………………..……(5)

Since Huffman coding is optimal, Eq. 5 will be violated if any bits are inserted. In the case of E < 1, from Table 1, we can find the relationship between the inserted bit location n belongs to A' (relative position within the codeword, left to right order) with Y=- Iog2 (1- E), for four original tree conditions. From Fig. 2, we can see that if u>=6 , we can obtain an unique solution of u from y . When 2<= u< 6 ,the number of solution is at most 2. Under this situation, we just keep these two candidates, and one of them can be used to successfully remove the inserted bit, leading to correct decoding of the bit stream.

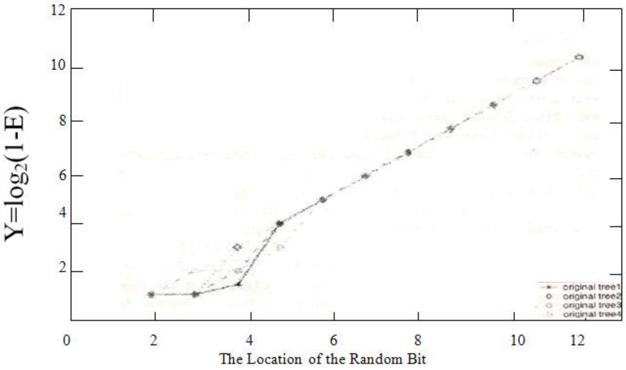

Figure 2. Relationship between inserted bit location u and Y

IV. HOW TO SELECT HUFFMAN TABLES

From the analysis of this area, we see that if the order that the m tables are used is fixed, the MHT schemes are inherently insecure. To overcome this drawback, a simple yet effective method is to use a secure stream cipher to generate the using order sequence [3], [4] and to avoid using the same sequence more than once- It should be pointed out that the security of this kind of improved scheme still heavily relies on the tables selected. In addition, to achieve a certain level of security, the number of tables is also closely related to the table selection. Therefore, a natural question arising is how to appropriately choose a small number of Huffinan tables to ensure a high level of security.

We observe that, in the MHT scheme, one symbol may be mapped to many different code words, and in turn, one codeword may also correspond to many different symbols. This many to many mapping results in the existence of mote than one symbol strings satisfying the same symbol distribution, which can be encoded to the same bit stream with a certain using order sequence. We call these symbol strings alias symbol strings. In order to evaluate the bidirectional mapping, we empirically define two quantities, namely, mapping diversity (MD) and codeword diversity (CD).

Consequently, when selecting the Huffman trees, our strategy is to assign more distinct code words to those symbols with high probability. In this way, we can simplify the stream cipher integrated scheme while ensuring a high level of security.

In this paper, we have not discussed the schemes with stream cipher generators. Generally speaking, the only relevant attack model in those cases is the known plain text attacks because the internal slate of the stream cipher is independent of the plaintext and the cipher text, making the flexibility of chosen-plaintext attack cannot be exploited. Currently, it is not practical to recover the secret keys of the stream cipher such as SHA-1, although recently breakthrough has been achieved [9], The security of these schemes, however, solely relies on the assumption that decoding a multiple Huffinan table encoded bit stream is computationally infeasible even if we have full knowledge about all the multiple tables, but only do not know the using order. To what extend is this assumption true is still an open question, and needs more mathematical analyses.

V. CONCLUSION

We have analyzed the security of the multimedia encryption Scheme using multiple Huffman table. The result of known plaintext attack shows that the multiple Huffman tables should be carefully selected to avoid the weak keys problem. Chosen plaintext attacks are then proposed to evaluate the basic MHT algorithm as well as the enhanced scheme with random bit insertion. Guided criteria for how to select the Huffman tables, we further suggest a simplified scheme to achieve a high level of security.

REFERENCES

[1] Data Encryption Std, FIPS PUBS 46-3,1999.

[2] Advanced Encryption Std., FIPS PUBS 197, 2001.

[3] Recommendation for Block Cipher Modes of Operation NIST Special Pub. 800-38A,2001.

[4] C.Wu and C.-C. Jf.Kuo, "Design of integrated multimedia compression and encryption systems," IEEE Trans. Multimedia,vol.7, no. 5, pp. 828-839, Oct. 2005.

[5] D. Xie and C.-C. J. Kuo, "Enhanced multiple Huffman table (MHT) encryption scheme using key hoping." in Proc.ISCAS, May 2004, vol. 5, pp. 568-571.

[6] "Multimedia data encryption via random rotation in partitioned bit stream," in Proc. ISCAS, May 2005, vol. 5, pp.5533-5536.

[7] I. Cheong, Y. Huang, and Y. Yung eta!., "An efficient encryption scheme for MPEG video," in Proc. ICCE, Jan. 2005, pp. 61-62.

[8] S. T. Klein, A. Bookstein, and S. Deerwester, "Storing Lest retrieval systems on CD-ROM: Compression and Encryption", ACM Trans. Information Systems, vol. 7, no.3, pp. 230-245, 1989.

[9] D. Gillman, M. Mohtashemi, and R Rivest, "On breaking a Huffman code," IEEE Trans. Inf. Theory, vol. 42, no. 3, pp. 972-976, Mar. 1996.

[10] C. Wu and C.-C. J. Kuo, "Efficient multimedia encryption via entropy codec design", SPIE international symposium on electronic imaging, San Jose, Jan 2001

[11] C. Wu and C.-C. J. Kuo, "Fast encryption methods for audiovisual data confidentiality", SPIE Photonics East-Symposiurr on Voice, Video, and Data Communications, Boston, Nov 2000.

[12] A.J.Menezes, P.C.van Oorschot and S A.Vanstone, "Handbook of Applied Cryptography", CRC Press, 1996

[13] D. Gilltnan, M Mohtashemi, and R. Rivest, "On breaking a Huffman code," IEEE Trans. Inf. Theory, vol. 42, no. 3, pp.972-976, Mar. 19%.

[14] W. Diffie and M. Hellman, "New directions in cryptography," IEEE Trans. Inf. Theory, vol. 22, no. IT-6, pp. 644-654, Jun. 1976.

Authors Profiles :

1) B.Prabhavathi, working as a assistant professor in AVR & SVR Engineering College, Nandyal. I have 1 Year Experience in teaching profession in CSE Department. I completed my M.Tech degree in Computer Science specialization in 2010. I am Interested in Networking & Data Mining.

2) P.D. Chidhambara Rao received the MCA degree from JNTU University, Kukatpalli, Hyderabad from the Department of Master of Computer Applications. He is a faculty member in the Department of MCA. My research interests are in areas of Computer Networks, interests are in computer vision, signal processing, and pattern recognition.

3).P.RajaSekhar,Asso.Professor,
SVPCET, R.V.S.Nagar,Puttur,
Chittur(Dt.),A.P.
Dept.of CSE,
My Interesting Area is
Networks & Ad-hoc routing protocols.
8 years of Teaching Experience.

4). Ishrath Jahan received the M. Tech degree in Computer Science and Engineering from JNTU Kakinada University from the Department of Master of Technology. She is a faculty member in the Department of M. Tech. My research interests are in areas of Computer Networks, Mobile Computing, DSP, MATLAB, Data Mining, Sensor Networks. She is working as CA faculty in Vijayawada. Delivered guest lecture in the subject DBMS & E- Commerce in NIST, Ibrahimpatnam for B.Tech students, Ibrahimpatnam. She worked as a guest lecturer in the subject Management Information System for MBA students in Nimra PG College.

5). G.Murali Mohan, Associate Professor, SVPCET, R.V.S.Nagar, Puttur, Chittur (Dt.), A.P.Department of CSE, My Interesting Area is Sensor Networks & Ad-hoc routing protocols. He has 7 years of Teaching Experience.