On the DCT Based Joint Fingerprint Embedding and Decryption Scheme

Xi Chen
Department of E-commerce, Nanjing University, Nanjing, China
Email: chenx@nju.edu.cn

Shiguo Lian
France Telecom R&D Beijing, Beijing, China
Email: shiguo.lian@ieee.org

Abstract—Till now, some Joint Fingerprinting and Decryption (JFD) schemes have been reported, while few works have been done to analyze their performances. In this paper, the security of the JFD scheme proposed by Kundur et al. is analyzed and improved. The analyses include the compliance with general compression codecs, the security against cryptographic attacks, the relationship between security and robustness, and the relationship between security and imperceptibility. Additionally, some means are proposed to improve the scheme's performances. The encryption in block based DCT is introduced to make it compliant with such compression codec as JPEG or MPEG2. The multi-key encryption is presented to improve the security against cryptographic attacks. The DC encryption is proposed to strengthen the perceptual security against ciphertext-only attacks. These analysis methods can also be used to evaluate some other JFD schemes, and are expected to provide valuable information to design JFD schemes.

Keywords—security analysis, fingerprinting, image encryption, joint fingerprinting and decryption (JFD), Digital Rights Management

I. INTRODUCTION

Secure multimedia distribution becomes more and more important with the wide application of multimedia data and the fast development of networks. Generally, encryption method [1,2] is used to protect media data’s confidentiality. Thus, only the authorized receiver can decrypt the data stream correctly. However, encryption method cannot solve the Super Distribution problem [3], that is, the decrypted copies can be distributed freely. To solve this problem, the fingerprint-based scheme [4,5,6,7] has been proposed. That is, for each customer, a unique watermark (fingerprint) is embedded into the multimedia program. Thus, each customer receives a different copy, and the fingerprint identifies the customer. This scheme can trace the illegal redistributors who send their copies to unauthorized customers.

One of the difficulties in fingerprint-based scheme is how to distribute the fingerprinted copies efficiently. Generally, there are three methods: embed fingerprint at the server end, in the router or at the receiver end. Straightforwardly, the server could embed a fingerprint in the media data and then send it to the according customer. However, considering that many customers may ask for the service at the same time, it is not practical for the server to send different copies to different customers efficiently. The second scheme, named WaterCasting, embeds watermark in the routers [6], which distributes the server’s loading to the sub-servers. However, the transmission protocol should be modified, which is not compliant with network protocol. Another scheme [7] embeds a fingerprint at the receiver end. Thus, the server sends only one media stream to all the customers. The potential difficulty is to keep the security of the operation at the customer end.

Embedding the fingerprint at the customer end may be practical if the security can be confirmed. Generally, media data are encrypted during transmission. Thus, at the receiver end, media data should be decrypted before displayed. If the fingerprint is embedded after the media data are decrypted, the decrypted media data may be leaked out before fingerprint embedding process. As an alternative, the joint fingerprint embedding and decryption (JFD) scheme [8,9] decrypts and embeds the media data at the same time, which avoids the leakage of the decrypted media data. Thus, different decryption key produces different copy, which is shown in Fig. 1. Here, only one encrypted stream is transmitted, and different copies are produced by different key. At the server end, the original data are encrypted under the control of KE. At the receiver end, the encrypted data are decrypted and fingerprinted with the JFD algorithm. In JFD, different decryption key $K_{D,j}$ ($j=0,1,\ldots,M-1$) (M is the number of the customers) generates the different decrypted and fingerprinted copy.
For JFD scheme, the security and robustness are emphasized. Till now, some methods have been proposed, e.g., Chamleon scheme [8], Kundur et al’s scheme [9], Lian et al's scheme [10] and Lemma et al's scheme [11]. Chamleon scheme is based on a stream cipher, which encrypts the media data with a stream cipher, decrypts and fingerprints the media data by modifying the LSB bits. This scheme is secure in cryptographic aspects [8], but is not robust to signal processing, such as recompression, adding noise, etc. For example, a slight noise can make the fingerprint in media data lost. This property makes it not practical. The scheme proposed by Kundur et al [10] encrypts media data at the server side by encrypting the variable-length code's index, and decrypts media data at the customer side by recovering code's index with both decryption and fingerprinting. This scheme is security against cryptographic attacks [8], while the robustness against general operations can not be confirmed. The scheme proposed by Lemma et al [11] encrypts media data at the server side by partial encryption, and decrypts media data at the customer side with a new key stream. It uses two different key streams for encryption and decryption respectively, which is similar with Chamleon method [8]. The scheme is robust against signal processing, which benefits from the adopted watermarking algorithms, while the security against cryptographic attacks can not be confirmed. Additionally, the transmission of key stream costs much time and space. Kundur et al’s scheme [9] is based on partial encryption, which confuses the sign bits of the DCT coefficients in encryption and decrypts only part of the sign bits in decryption. The position of the left sign bits determines the fingerprint. This scheme is often robust to some signal processing operations. Compared with previous schemes, Kundur et al's scheme is more suitable for compressed data. However, considering that most of the quantized coefficients are zeros, the method’s security should be further discussed. Additionally, the number of the coefficients to be encrypted is in close relation with the robustness and imperceptibility.

Recently, some other JFD schemes are reported. For example, the method in [17] modulates the media content with noise sequences at the sender side, and de-modulates the content selectively at the receiver side. The method proposed in [18] encrypts the significant part of media content, while embedding unique information into the non-significant part of media content. In [19], the commutative watermarking and encryption operations are adopted to encrypt and mark media contents before media distribution. These schemes’ performances have been tested and analyzed. However, there are still not general principles for designing JDF schemes.

Till now, few works have been done to discuss JFD schemes’ performances thoroughly. The work in [20] considers a joint audio fingerprinting and decryption scheme. In this paper, taking Kundur et al's scheme for example, the performances of the JFD scheme are investigated and analyzed. We try to give a detailed discussion on its security, point out the relationship between its security and its robustness, imperceptibility, and propose some means to improve its performances. It is expected to provide valuable advices to design a secure JFD scheme. This work's initial version has been put on the e-print archive [24].

The rest of the paper is arranged as follows. In Section 2, Kundur et al’s scheme is briefly introduced. The scheme's effect on compression is evaluated in Section 3, and its cryptographic security is analyzed in Section 4. In Section 5 and Section 6, the relationship between security and robustness, and the one between security and imperceptibility are investigated in detail, respectively. The perceptual security of the partial encryption scheme is analyzed in Section 7. In Section 8, some means are proposed to improve the scheme’s performances. Finally, conclusions are drawn, and future work is given in Section 9.

II. KUNDUR ET AL.’S JFD SCHEME AND ITS ALTERNATIVE IMPLEMENTATION

A. Introduction to the Proposed JFD Scheme

Kundur et al proposed the joint fingerprint embedding and decryption scheme based on partial encryption [9]. In encryption, the whole image is transformed by DCT, DCT coefficients' signs are partitioned into n subsets, each of which is composed of Mb signs, and all the subsets are encrypted under the control of KE. In decryption, only p (0<p<n) subsets of signs are decrypted, while the other n-p subsets are left unchanged. Which subsets to be decrypted are determined by the decryption key Kd,j (j=0,1,....,M-1) (M is the number of the customers). Different customer has a different decryption key, and different decryption key produces a different
copy. In each copy, the positions of the undecrypted subsets determine the image copy’s uniqueness. The encryption and decryption processes are shown in Fig. 2. In sign partitioning, the subsets can be overlapped. The following parameters are recommended by the authors, i.e., \( n=25, \ p=19, \ M_b=200 \), which can obtain the encrypted image with PSNR=22dB and the fingerprinted image with PSNR=34dB.

B. Alternative Implementation of Kundur et al.’s JFD Scheme

Considering that the coefficients in the low frequency of DCT domain are sensitive to images’ intelligibility, they should be decrypted at the receiver end, which has not been considered in [9]. In the following content, we tend to leave only some coefficients in the middle or high frequency undecrypted. For convenience, some parameters are set: \( N, N_T \) and \( N_{\text{nonzero}} \). Among them, \( N \) denotes the total number of the DCT coefficient, \( N_{\text{nonzero}} (0<N_{\text{nonzero}}<N) \) denotes the number of nonzero coefficients that are actually encrypted at the server end while \( N_T (0<N_T\leq N_{\text{nonzero}}) \) denotes the number of coefficients that are decrypted at the receiver end. Thus, the \( N \) coefficients are operated according to the following steps (as shown in Fig. 3):

i) Zigzag scan: \( N \) coefficients are first ordered in zigzag mode \( x_0x_1...x_{N-1} \);

ii) Sign encryption: The first \( N_{\text{nonzero}} \) coefficients \( x_0x_1...x_{N_{\text{nonzero}}-1} \) are encrypted;

iii) Partial decryption: The first \( N_T \) coefficients \( x_0x_1...x_{N_T-1} \) are decrypted, while the left \( N_{\text{nonzero}}-N_T \) coefficients are partially decrypted;

iv) Inv-zigzag scan: The processed coefficients are set back to the DCT block in inv-zigzag order.

III. NON-COMPLIANT WITH COMPRESSION

In this scheme, the whole image is transformed by DCT, and then, the signs are encrypted. Considering that most of the existing image or video compression methods apply DCT to the blocks with the size of 8x8 or smaller, this scheme can not be used in compression domain.

<table>
<thead>
<tr>
<th>Images</th>
<th>Changes on JPEG Compression (%)</th>
<th>Changes on MPEG2 Compression (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena (512x512)</td>
<td>21 Foreman (CIF) 14</td>
<td></td>
</tr>
<tr>
<td>Airplane (256x256)</td>
<td>18 Mobile (QCIF) 17</td>
<td></td>
</tr>
<tr>
<td>Boats (512x512)</td>
<td>24 Football (CIF) 20</td>
<td></td>
</tr>
<tr>
<td>Baboon (256x256)</td>
<td>26 Salesman (QCIF) 16</td>
<td></td>
</tr>
</tbody>
</table>

That is, the sign change of the whole image's DCT coefficients will lead to great changes to image block's DCT coefficients, which change the statistic distribution of DCT coefficients, and thus affect the followed entropy coder's compression efficiency. If using the encryption scheme to encrypt the image or video before compression, e.g., JPEG [16] or MPEG2, the scheme will affect the compression efficiency, as shown in Table 1. Here, the recommended parameters (\( n=25, \ p=19, \ M_b=200 \)) are adopted, and the changes are tested by the ratio between the increased data size and the unencrypted data size.

IV. THE CRYPTOGRAPHIC SECURITY

This JFD scheme is first an encryption scheme, and also a watermarking or fingerprinting scheme. As an encryption scheme, the cryptographic security should be analyzed. Different from traditional encryption schemes, two attack cases arise in Kundur’s scheme. One case is that the unauthorized customer tries to obtain the plain-media, which is named unauthorized attack by us. The other case is that the authorized customer tries to obtain another plain-media, which is named authorized
attack by us. The security against both the attacks is presented as follows, respectively.

**A. Security against Unauthorized Attack**

In unauthorized attacks, the unauthorized customers have no keys to decrypt the media data. To get the plain-media is attractive to them. The direct method is the brute-force method. Generally, the brute-force space of a cryptosystem is the key space. However, in Kundur’s scheme, the case is different. That is because the one encryption key corresponds to several decryption keys. Set the encryption space of KE be S (S is the key’s space, for example S=2^64 if the key is of 64-bit length). The relation between the encryption key and the decryption key is shown in Fig. 4(a). Here, for each encryption key KE, M decryption keys KD,0, KD,1, ..., KD,M-1 can be used to decrypt the media data into an intelligible copy. Thus, the brute-force space is reduced to

\[
S_{\text{un}}(M) = S - M + 1. \tag{1}
\]

That is, an unauthorized customer needs only to try S-M+1 times but not S times before he can obtain the plain-media. Here, the space is a function of the customer number M under certain S. The relation is shown in Fig. 4(b) where S=2^64. As can be seen, under certain key space S, the brute-force space decreases with the rise of the customer number M. Thus, in order to keep secure, the key space S should be big enough, otherwise, the customer number M should be keep small enough.

\[
S_{\text{sign}} = 2^{N_{\text{nonzero}}}. \tag{3}
\]

**B. Quantization and \( N_{\text{nonzero}} \)**

In quantization, the quantization factor q determines the number of nonzero coefficients. Generally, the bigger the quantization factor q is, the smaller \( N_{\text{nonzero}} \) is. That is, \( N_{\text{nonzero}} \) and q satisfy the following condition.

\[
N_{\text{nonzero}} = o\left(\frac{1}{q}\right) \tag{4}
\]
Fig. 5 shows the statistical result on Lena. Here, the quantization table in JPEG [16] is adopted, and q ranges from 0.1 to 3. As can be seen, the relation between N_{nonzero} and q is a decreasing function.

C. Security and Quantization

The relation between security and quantization satisfies

\[ S_{\text{sign}} = 2^{\frac{q}{1+q}}. \]  

(5)

Thus, the encryption space decreases with the rise of the quantization factor. The relation is shown in Fig. 6. That is, for the media data with high quality (small q), the encryption system is more secure. Otherwise, if the media data are greatly quantized, few nonzero coefficients are left, and the scheme is of lower security.

![Fig. 6. Encryption Space against Quantization](image)

VI. THE RELATIONSHIP BETWEEN SECURITY AND IMPERCEPTIBILITY

In Kundur’s scheme, not all the signs of the coefficients can be left unencrypted. In order to keep the fingerprinted media data imperceptible, only the coefficients that are not sensitive to perceptual quality can be left. Thus, except zero coefficients and the sensitive ones, the remaining coefficients can be used to fingerprint the media data. The number of the remaining coefficients determines the number of the customers the scheme can support. Thus, there is relation between security and imperceptibility, which is presented as follows.

A. Imperceptibility and Threshold Coefficient

According to the distribution property of DCT coefficients, the sensitivity decreases with the rise of the coefficient’s frequency, as shown in Fig. 7. Thus, the threshold coefficient N_T is defined, which denotes the threshold position, from which to higher frequency, the coefficients can be fingerprinted. Generally, the imperceptibility I and threshold coefficient N_T satisfy the following condition

\[ N_T = o(I). \]

(6)

That is, the higher the imperceptibility is required, the bigger the threshold coefficient N_T should be. Fig. 7 shows the sensitivity of different coefficient, the relation between the threshold coefficient N_T and the number of nonzero coefficients N_{nonzero}. As can be seen, the coefficient’s sensitivity decreases with the rise of the no. of coefficient. In order to keep the PSNR no lower than 55, the threshold N_T should no smaller than 7. Additionally, the condition should be satisfied: N_T < N_{nonzero}.

![Fig. 7. Coefficient Sensitivity in DCT Transformation](image)

B. Security and Threshold Coefficient

In Kundur’s scheme, the first N_T coefficients should be decrypted completely, while the remained N_{nonzero}-N_T ones are fingerprinted. Thus, the number of customers is determined by

\[ M = 2^{N_{\text{nonzero}}-N_T}. \]

(7)

The bigger N_T is, the fewer the customers can be supported. Thus, according to (1), the brute-force space is

\[ f_{\text{brute}}(M) = S - 2^{N_{\text{nonzero}}-N_T} + 1. \]

(8)

That is, the bigger N_T is, the larger the brute-force space is.

C. Imperceptibility and Security

According to the above analysis, the relation between imperceptibility and security is

\[ f_{\text{brute}}(M) = S - 2^{N_{\text{nonzero}}-I} + 1 \]

(9)

It shows that the higher the imperceptibility I is required, the larger the brute-force space is.

VII. THE PERCEPTUAL SECURITY

For media encryption, perceptual security [1][2] is often considered besides cryptographic security, which means that the encrypted media data should be unintelligible. Otherwise, if the encrypted image is still intelligible, the image's quality can be improved by using the image's statistic model [1]. In Kundur’s scheme, the DCT coefficients in low frequency band are preferred to be encrypted, and the more the coefficients are encrypted,
the higher the perception security is. Fig. 8(a) shows the relation between the number of encrypted coefficients Nen and the media quality (Cameraman, q=0.5). Thus, in sign encryption, the Nnonzero nonzero coefficients are all encrypted. However, in Kundur et al’s scheme, only DCT coefficients’ signs are encrypted, while the coefficients’ amplitudes are kept unchanged, which makes the encrypted image in Fig. 8(b) still intelligible. Here, the encryption parameters (n=25, p=19, M=200) recommended in [9] are used. According to the attack method proposed in [1], we can easily recover the image’s quality even without the decryption key and obtains the image in Fig. 8(c). As can be seen, the recovered image has higher quality that can even be improved by filtering methods. Therefore, Kundur et al’s scheme is not secure in perception and can be easily hacked by existing methods.

![Image](image-url)

Figure 8. Perceptual Security of Sign Encryption

VIII. MEANS TO IMPROVE THE SCHEME’S PERFORMANCES

According to the above analysis, the security of Kundur’s scheme is in close relation with the scheme’s robustness and imperceptibility. To keep high robustness, the number of nonzero coefficient Nnonzero should be decreased, although it reduces the security in some extent. To keep good imperceptibility, the threshold coefficient NT should be increased, although it reduces the number of customers that can be supported. In order to improve the scheme’s performances, several means are presented as follows.

A. Data Encryption in Block Based DCT

To adapt with compression methods, e.g., JPEG or MPEG2, DCT block encryption is proposed, which is different from the whole image encryption in Kundur et al’s scheme. That is, the DCT signs of DCT blocks in JPEG or MPEG2 are grouped into subsets and encrypted. Considering that the coefficients in the low frequency of DCT blocks are sensitive to images’ intelligibility, they should be decrypted at the receiver side. Thus, we leave only some coefficients in the middle or high frequency undeeprecated. For convenience, we set N the total number of encrypted DCT coefficient in each block, and NT (0<NT≤N) coefficients should be decrypted at the receiver side.

B. Multi-Key Encryption

As is mentioned in Section 3, the brute-force space decreases with the rise of the number of customers that can be supported. To enlarge the brute-force space is a solution to this problem. That is, the reduced brute-force space still satisfies the requirement of practical applications. Thus, multi-key encryption can be used here: media data are partitioned into H parts, and each part is encrypted with a different key. In this case, the brute-force space is

\[
\begin{align*}
\text{f}_{\text{un}}(M) &= S^H - M + 1 \\
\text{f}_{\text{un}}(M) &= S^H - M \\
\end{align*}
\]

Compared with the previous one, the space is greatly enlarged, and the number of the customer does little effect on the scheme’s security.

C. Selection of NT and Nnonzero

The number of Nnonzero is in close relation with the security and robustness. Generally, it is computed by setting a suitable quantization factor. According to the required robustness (the supported compression ratio in JPEG/MPEG), the maximal quantization factor can be obtained, which is used to generate Nnonzero. Generally, under q=0.5, the average value is Nnonzero=16. Differently, the threshold coefficient NT determines both the imperceptibility and the number of customers. For NT, a tradeoff should be determined between the imperceptibility and the number of customers.
coefficients in low frequency causes great blurs to images. low frequency after DCT transformation, encrypting the block artifact degrades image quality greatly. Thus, the maximal number of customer that the scheme can support is $M=2^L$. Additionally, it is practical to enlarge the number by multi-key method. That is, the image is partitioned into $L$ parts, and each part is decrypted with different mode. Thus, the maximal customer number is changed into $M=2^L$. The encryption operation should be applied following quantization process, in order to make the image decrypted correctly. Fig. 10 shows the encrypted image and decrypted image, respectively. Here, the encrypted image (a) is unintelligible, which keeps secure against the recovering method [17][21][22][23]. Thus, the improved scheme obtains higher perceptual security. Additionally, the decrypted image with the fingerprinted embedded keeps high imperceptibility.

![Figure 10. Perceptual Security and Imperceptibility of the Improved Scheme](image)

**IX. CONCLUSIONS AND FUTURE WORK**

In this paper, Kundur et al’s Joint Fingerprinting and Decryption scheme is evaluated. This scheme encrypts the DCT signs after transforming the whole image with a DCT, which is not compliant with existing compression methods. The scheme's cryptographic security decreases with the rise of the number of supported customers. Additionally, the low perceptual security makes the scheme fragile to some ciphertext-only attacks that adopt the image's statistical model to recover the original image. Some means are proposed to improve the scheme's security and make it compliant with compression processes, such as data encryption in block based DCT, multi-key encryption, and DC encryption. The evaluation on the scheme's other performances, e.g., collusion-resistance or computational efficiency, is the future work. Additionally, the analysis method can be used to evaluate some other Joint Fingerprinting and Decryption schemes.

**ACKNOWLEDGEMENTS**

The work was supported in part by the National Natural Science Foundation of China under Grant No. 70901039 and 70732002, National Postdoctoral Science Foundation of China under Grant No. 20090450144, Jiangsu Postdoctoral Science Foundation under Grant No. 0901104C and France Telecom R&D Beijing.

**REFERENCES**


Xi Chen received Ph.D. from Nanjing University of Science and Technology, China. He is an associate professor in Management School of Nanjing University and a visiting scholar of University of Washington. He authors more than 40 refereed journal/international conference papers and chapters. He authored the book "The theory of enterprise resource planning". He got the “Best Paper Awards” from Chinese Academy of System Simulation. He is the project chief of National Science Foundation and National Ministry of Education Science Foundation. His research interests include service science, E-Commerce security, information management and data authentication. He is a member of some Technical Committees, the technical committee chair and organization committee chair of refereed conferences/workshops, the peer review expert of NSFC. He is on the editor board of several international journals, and reviewer of some refereed international journals and conferences.

© 2010 ACADEMY PUBLISHER

Shiguo Lian got his Ph.D. from Nanjing University of Science and Technology, China. He was a research assistant in City University of Hong Kong in 2004. Since July 2005, he has been a Research Scientist with France Telecom R&D (Orange Labs) Beijing. He is the author or co-author of more than 80 refereed international journal and conference papers covering topics of secure multimedia communication, intelligent multimedia services, and ubiquitous communication. He is a member of IEEE Communications & Information Security Technical Committee, IEEE Multimedia Communications Technical Committee, and IEEE Technical Committee on Nonlinear Circuits and Systems. He is on the editor board of several international journals. He is the guest editor of more than 10 international journals. He is in the organization committee or the TPC member of refereed conferences. He is also the reviewer of refereed international magazines and journals.