コンテンツ保護機能を備えたインターネット生放送システムの実現可能性の研究

黄　亮錦†　　森村　吉貴‡　　上原　哲太郎§　　上田　浩¶　　津田　侑†

†京都大学　大学院情報学研究科
606-8501 京都市左京区吉田本町
{lj.hu, tsuda}@ipe.media.kyoto-u.ac.jp

‡京都大学　物質-細胞統合システム拠点
606-8501 京都市左京区吉田牛ノ宮町
ymorimura@icems.kyoto-u.ac.jp

§NPO 法人　情報セキュリティ研究所
646-0011 和歌山県田辺市新庄町 3353-9 Big-U
uehara@tetsutaro.jp

¶京都大学　学術情報メディアセンター
606-8501 京都市左京区吉田二本松町
uep@media.kyoto-u.ac.jp

あらまし USTREAM のようなインターネット生放送の Web サイトの人気や配信技術の向上や高速インターネット環境の成熟により, テレビ放送に取って代わるコンテンツ課金型のインターネット生放送の需要が高まってきている. コンテンツ課金型のインターネット生放送を展開する場合は課金したユーザのみがインターネット生放送を視聴できるような仕組みが必要となる. そこで我々は課金したユーザのみがその映像を視聴できるようにするために, 映像に対して JFD（Joint Fingerprinting and Decyption）と呼ばれる手法を用いた暗号化と電子指紋の埋込みを同時に行う仕組みを持ったインターネット生放送を開発した. そして, 開発したインターネット生放送の実現可能性を評価した.

A Feasibility Study of an Internet Live Broadcasting System with Contents Protection

Liangjin Huang†　　Yoshitaka Morimura‡　　Tetsutaro Uehara§　　Hiroshi Ueda¶
Yu Tsuda†

†Graduate School of Informatics, Kyoto University
Yoshida Honmachi, Sakyo-ku, Kyoto, 606-8501, JAPAN

‡iCeMS, Kyoto University
Yoshida Ushinomiya-cho, Sakyo-ku, Kyoto, 606-8501, JAPAN

§The Research Institute of Information Security
Big-U, 3353-9 Shinjo-cho, Tanabe-shi, Wakayama, 646-0011, JAPAN

¶Academic Center of Computing and Media Studies, Kyoto University
Yoshida Nihonmatsu-cho, Sakyo-ku, Kyoto, 606-8501, JAPAN

Abstract With the popularity of live broadcasting websites such as USTREAM, the maturity of live broadcasting technology, and the improvement of network condition, more and more multimedia service providers are using Internet to do the pay-per-view live broadcasting business instead of traditional TV platform. It becomes an issue which is how to provide an integrated protection scheme to the multimedia contents in the pay-per-view business. For our research, we aim to develop an Internet live broadcasting system with contents protection to solve this issue. We use Home Page cryptosystem to encrypt the contents and use JFD (Joint Fingerprinting and Decryption) method to embed the fingerprint. And it proved feasible according to the evaluation to the system.

1 Introduction

YouTube[1] made it possible for people with an Internet connection to share their lives not only by words and pictures but by videos. It turned the video sharing into one of the most important parts of Internet culture. However, YouTube can't meet the demand of some people who want to share with others immediately. For them, the Internet live broadcasting website is born, such as USTREAM[2]. It allows people to share with others in real-time, and it has a better interaction capability.

While the Internet live broadcasting technology becomes more and more mature, the network condition becomes much better than ever, for example, wider network bandwidth, less network congestion and so on, the quality of multimedia via Internet live broadcasting could already compare beauty with television. Besides, because of its costs is much less than the television, more and more organizations and companies realized the potential of using Internet live broadcasting system to do pay-per-view business and they've already developed kinds of interrelated businesses.

In the Internet pay-per-view business, the multimedia contents is valuable, we should take measures to protect them. The protection to the multimedia contents is also the protection to the payers' interests and the protection to the service providers' copyrights. The key point of contents protection here is to make sure only the payers have the rights to view the multimedia contents. Such a guarantee contains two aspects:

1. The multimedia contents is invisible to the unauthorized persons;
2. Since the payer gets the right to view the contents, we should inhibit the situation which the payer leaks the contents out.

To meet such a guarantee, we use broadcast encryption to encrypt the multimedia contents. However, broadcast encryption can't provide protection after the contents are decrypted, we use fingerprinting embedding method to deal with this problem. Furthermore, in case that the decrypted contents without fingerprint embedded are extracted by traitors who are among the payers but wants to leak the protected contents out, we should ensure that the decryption and the fingerprint embedding are processed in the same time.

2 Previous Work

Hou *et al.* proposed a contents protection scheme based on integrating anti-collusion code and Home Page public-key cryptosystem[1]. Fir-stly, they generate the public-key pair for encryption and descryption. They encrypt the multimedia contents by encrypting DCT coefficients of I-frame, where the I-frame is divided into plenty of 8*8 pixels' blocks with DCT transformed, and quantized. At the payer end, payer decrypts the encrypted contents by using the key from the service provider; at the same time, payer's personal fingerprinting is embedded into the decrypted contents. When the multimedia contents are leaked out, it would become very easy to find the source by detecting the fingerprinting.

Morimura *et al.* implemented the JFD based contents protection scheme[2] and did rounded evaluation proved that such a scheme was feasible on the bandwidth consumption and process speed. However, because the linux kernel function was called in his implementation, the high platform dependency is the weak point.

For this paper, we aimed to develop a live broadcasting system with contents protection. Differ from previous work, we pay much more attention to the clients who would use this system. We've talked about this system's object

[1] http://www.youtube.com/

[2] http://www.ustream.tv/

is for Internet pay-per-view business use, so figuring out what kinds of pay-per-view business suits for this system with contents protection is also an important part for our research. What's more, we should make sure such a system is easy and comfortable for clients to use, which means it should have a better performance on simple operation capability, image quality, program efficiency and so on. For this, we make a feasibility study on the system by evaluating the speed of encryption and decryption, image quality after encryption and decryption and so on.

3 The Internet Live Broadcasting System with Contents Protection

3.1 System Overview

Figure 1 shows the system's whole structure.

This system is divided into two parts, server end and client end. For the platform independency, both server end and client end are implemented as ActiveX controls[3]. For the server end, it consists of the key-generation module, multimedia contents capturing previewing module, contents encryption module and network broadcasting module; for the client end, it consists of network receiving module to receiving the multimedia contents data from the server end, decryption and fingerprinting embedding module (which is the JFD module), contents playback module and the recording module. The detail information of the system's running process and each module would be introduced later.

3.2 Applicable Businesses to the System

Before doing a close introduction to the live broadcasting system with contents protection, it's necessary to figure out that the applicable contents to this system should meet the needs of real-time, value and easy-getting.

1. The need of real-time means that the multimedia contents come from broadcaster needs to be sent to clients in real-time. For instance, a football match can make people excited due to the match's nondeterminacy, that's why people prefer watching the live broadcast to the recorded broadcast.

2. The need of value means that the multimedia contents come from the broadcaster should be attractive enough for people to purchase. People won't pay for the movie which is boring enough.

3. The need of easy-getting means the multimedia contents come from the broadcaster should be easy to get just by some video and audio capture devices.

A football match may meet the need of real-time and value, but due to the huge playground and the quick ball movement, it doesn't meet the need of easy-getting.

Let's take online classroom service into consideration. Differ from traditional classroom education, online classroom overcomes the issue of geographic limitation. Due to the real-time capacity, it's an interactive education so it's better than audio-visual product education. The contents in the online education also meet the need of value. Furthermore, it meets the need of easy-getting because what the online classroom needs is just a camera, a microphone, a computer, an Internet access, a quiet room and a talented teacher.

3.3 Simulation of System's Execution

In order to get a better understanding to the live broadcasting system with contents protection, we use a situation simulation to exy

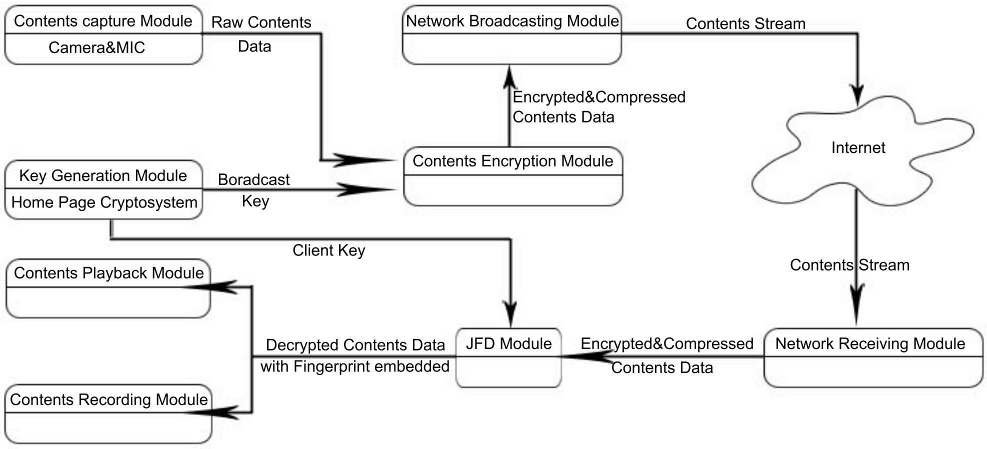

Figure 1: System's Architecture

plain the system's execution progress. Here, we choose the pay-per-view online classroom service as the simulative business.

1. A student named Alen purchased the online classroom service.
2. System modules testing. Before the class begins, each system's module should be tested making sure they are working well.
3. Teacher Bill begins to give online lessons. Firstly, Bill starts the key-generation module then generates a pair of keys. Public key which is the broadcast key, is used to encrypt the multimedia contents; private key which is the client key, is used to decrypt the encrypted multimedia contents. Bill sends the client key to Alen via a secure channel. Then Bill starts the multimedia contents capturing module to capture his lecture data via the video/audio capture devices. The captured lecture multimedia contents data is passed to the contents encryption module and encrypted by this module. Finally the network broadcasting module broadcasts the encrypted contents data flow to the Internet.
4. At the same time, on student Alen's end, system's client part receives the encrypted contents data from the Internet by the network receiving module, then passes the encrypted data to the JFD module. In the JFD module, decryption and fingerprinting embedding to the encrypted contents data are progressed simultaneously. Decrypted contents data with fingerprinting embedded is sent to contents playback module so that Alen could view the live lecture. Furthermore, this system provides the recording function for Alen to record the lecture for review.
5. The lecture records are only for student themselves to review. If someone breaks the agreement, gives publicity to the lecture contents on the Internet illegally, this system could extract the fingerprint from the contents then find out the people who leak the lecture contents out.

3.4 System Composition Modules

This system is divided into server end and client end. The detailed description of server end's modules is as follows.

1. Key-generation module: we've already known that we need a pair of keys to encrypt and decrypt the multimedia contents. Here is the Home Page cryptosystem[4] which is used to generate the keys. By comparing

	RSA	HP
Category	Trapdoor One-Way Function Cryptosystem	Knapsack Cryptosystem
O	Exponentiation	Multiply-accumulate
Public Key	e, n	k$_{s_1 1}, k_{s_1 2} ... k_{s_1 I}, k_{s_2 1} ... k_{s_j i} ... k_{s_J I} (J * I)$
Private Key	d	d$_i, V_i, P, w(I)$
Encryption	$C = P^e mod\ n$	$C = m_1 * k_{s_i 1} + m_2 * k_{s_i 2} + ... + m_I * k_{s_i I}$
Decryption	$P = C^d mod\ n$	$L = w^{-1} * C(mod\ P)$ $m_i = L * V^{-1} (mod\ d_i)$
Security Foundation	Difficulty of Prime Decomposition	Huge Quantity of Public Key's Combination

Table 1: The Comparison between RSA and HP

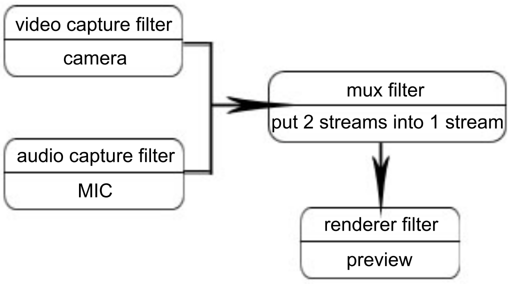

Figure 2: Multimedia contents capturing previewing module

with the RSA in Table 1, we could have a general idea to the Home Page cryptosystem.

2. Multimedia contents capturing previewing module: this module is implemented by using Microsoft DirectShow API[5]. DirectShow API is a media-streaming architecture for Microsoft Windows. With it, the applications can perform high-quality video and audio playback or capture. Figure 2 shows the structure of capturing preview module, which captures the video via camera and the audio via MIC, then previews the AV contents on screen in real time.

3. Contents encryption module: In this module, we use the public key generated by Home Page cryptosystem to encrypt the contents come from contents encryption module. We modify the MPEG4 encoder in ffmpeg's video codec library[6] to equip the MPEG4 encoder with encryption function: After the quantification of DCT coefficients, we choose the middle-frequency coefficients as the plaintext M to encrypt, ciphertext C is saved into the user data filed which is a part of frame header. The original middle-frequency coefficients are set to zero so that the encrypted video is presented as a mess. Figure 3 shows the encryption process clearly.

4. Network broadcasting module: In this module, we do the encrypted contents data transmission by using JRTPLIB[7] JRTPLIB is an object-oriented library which aims to help developers in using the Realtime Transport Protocol to transport the multimedia data.

On the client end:

1. Network receiving module: The same as the Network broadcasting module, we receive the multimedia stream comes from the server end by using JRTPLIB.

2. JFD module: This module is implemented based on the MPEG4 decoder in ffmpeg video codec library. The MPEG4 decoder's process is almost the MPEG4 encoder's opposite. We do the decrypt operation before the inverse quantization. First, we get

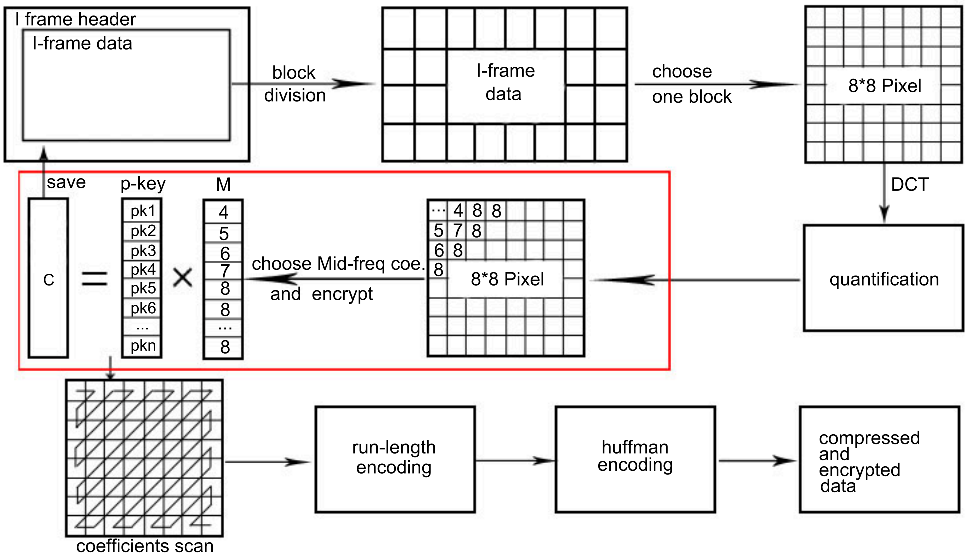

Figure 3: Contents encryption module

the ciphertext from the frame header one by one, each ciphertext is related to an 8*8 pixel block. In order to embed fingerprint while decrypting, the original private key is modified, dummy keys replace some of factors in original private key, and the decrypted value of the corresponding mi where dummy keys are replaced would be 0. With the dummy keys variable quantities and position, the fingerprint is embedded. According to the formula below,

$$L = w^{-1} * C(mod\ P) \qquad (1)$$

$$m_i = L * V^{-1}(mod\ d_i) \qquad (2)$$

(L stands for intermediate ciphertext, C stands for ciphertext, m_i stands for decrypted message, P, w, V, d_i, stands for private key) we get the decrypted plaintext which is the middle-frequency coefficients, we put them back to their original position in the block, finally the JFD operation is finished.

3. Contents playback module and the recording module: Both of these two modules are implemented by DirectShow API. Contents playback module is in charge of the playback of decrypted contents with fingerprint embedded; recording module is in charge of saving the data on the disk.

4 The System's Evaluation

4.1 Overview of Evaluation

For such a live broadcasting system with contents protection, the important evaluation factors are including: the efficiency of encryption and decryption, the security and real time capability of data, the image quality and so on. For the data security, because of the application of HP cryptosystem and JFD scheme, the probability of illegal decryption to the encrypted data and falsification to the digital fingerprint is proved very low in [1]. To our research, user experience comes first, we pay much more attention to the image quality, efficiency of encryption and efficiency of decryption.

4.2 Evaluation Items

1. Qscale

   Qscale is a very important parameter in MPEG4 codec. The value of Qscale determines the video's image quality and video

No.	Qscale	video size	image quality
1	0.01	819KB	
2	1	789KB	
3	5	216KB	
4	10	97KB	
5	20	54KB	

Table 2: The Relationship among Qscale, Image Quality and Video Size

size. The higher Qscale value, the prettier imapge quality, though the bigger video size. Table 2 shows the trade-off relationship among Qscale, image quality and video size. Finally we choose the value 5 as the Qscale's value. The five videos have the same data in duration (5 seconds), resolution (320*240), GOP number which is 12 frames (IPPP PPPP PPPP), and FPS (20fps), they are all encoded by the same MPEG4 encoder. Moreover, the video contents are almost the same. Finally we choose the No.3 as the Qscale value, for it has a better balance between image quality and video size.

2. Image Quality

   The encryption to the original multimedia contents is to make the contents be unrecognized and insignificant to these unauthorized people who don't purchase the pay-per-view services. Figure 4 is the one of encrypted video's frames. Likewise, we need the decrypted multimedia contents to be clearly visible. Figure 5 is the example of decrypted video's frames with fingerprint embedded.

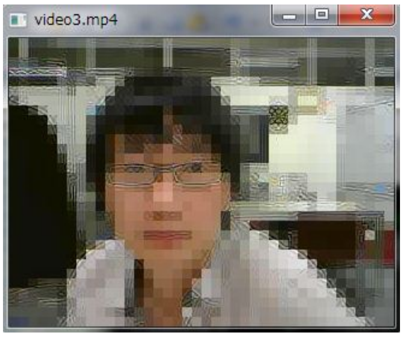

Figure 4: Encrypted Video's Frame

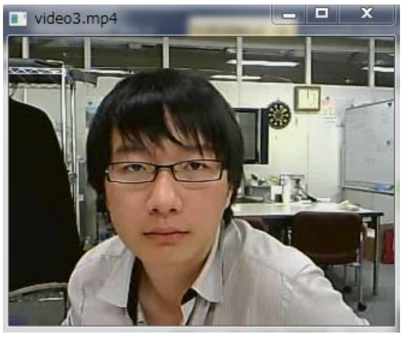

Figure 5: Decrypted Video's Frames with Fingerprint Embedded

5 Conclusion and Future Work

By using the Home Page cryptosystem and JFD scheme, we develop a live broadcasting system with contents protection, providing rounded protection to the multimedia contents not only during the encryption but also after decryption. In addition, we pay much more attention to the clients, for making the clients who use this system feel easy and comfortable, we develop this system to make it get a better performance on simple operation capability, image quality and so on. Besides, due to not all kinds of pay-per-view businesses are suit for this system, we also define the three standards to evaluate whether a business is necessary to use this system.

we evaluate user experiences of our system on having examinee use the system via Internet, such as usability of the system, feeling of viewing videos on real time and so on. Furthermore, we test the percentage of delays caused by encryption and decryption. Moreover, with the clients increasing, we need to find an efficient way to manage the key pairs.

Reference

[1] Shuhui Hou, Tetsutaro Uehara, T Satoh, Yoshitaka Morimura, and Michihiko Minoh, “Integrating Fingerprint with Cryptosystem for Internet-based Live Pay-TV System,” Wiley Journal Security and Communication Networks, Vol.1, No.6, pp.461-472, 2008.

[2] Yoshitaka Morimura, Tetsutaro Uehara, Shuhui Hou, “The Construction and Evaluation to the JFD System which for the Internet Broadcasting”, Trans. of the IEICE. Vol.J94-B No.10, pp.1427-1439, 2001 (Japanese).

[3] ActiveX controls
`http://msdn.microsoft.com/en-us/library/aa751968(v=vs.85).aspx/`

[4] Masao Kasahara, Yasuyuki Murakami, “New public-key cryptosystems” Technical Report of IEICE (ISEC), Vol.99, No.208, pp.33-99, 1999 (Japanese).

[5] Microsoft DirectShow API
`http://msdn.microsoft.com/en-us/library/windows/desktop/dd375454(v=vs.85).aspx/`

[6] FFmpeg
`http://ffmpeg.org/`

[7] JRTPLIB
`http://research.edm.uhasselt.be/j̃ori/page/index.php?n=CS.Jrtplib/`