Key Derivation Scheme for Hierarchical Access Control to Multimedia Content

Shoko Imaizumi, Naokazu Aoki, Hiroyuki Kobayashi
Graduate School of Advanced Integration Science
Chiba University
Chiba, Japan
Email: imaizumi@chiba-u.jp,
{aoki, kobahiro}@faculty.chiba-u.jp

Masaaki Fujiyoshi, Hitoshi Kiya
Dept. of Information and Communication Systems
Tokyo Metropolitan University
Tokyo, Japan
Email: fujiyoshi-masaaki@tmu.ac.jp,
kiya@sd.tmu.ac.jp

***Abstract*—This paper proposes an efficient key derivation scheme introducing a recursive hash chain for hierarchical access control. The proposed scheme has become more scalable and more flexible access control than the conventional schemes. The scheme uses both ordinary hash chains and a recursive hash chain to decrease the number of managed keys to one. In addition to reduction of managed keys, regardless of the number of the controlled media and those depths, the number of delivered keys for each user in the proposed scheme is equal to or less than that in the conventional schemes. The single managed key is not delivered to any user, providing security against key leakage. The proposed scheme also has collusion attack resilience.**

I. Introduction

With the continuing growth in network technology, the exchange of digital images and digital audio as well as digital text data has become very common. Since digital content is easily duplicated and re-distributed, protecting copyrights and privacy is an important issue. For the protection of digital content, *access control* based on naïve encryption (encrypting the whole content) [1] or media-aware encryption [2]–[6] has been studied widely.

A simple and straightforward way to realize versatile access control for multimedia content, consisting of several kinds of media to which several entities belong, is encrypting each entity individually. This approach, however, has to manage a large number of keys, given the large number of entities in multimedia content.

Scalable access control schemes have been proposed [2]–[6] for JPEG 2000 [7] coded images and/or MPEG-4 fine granularity scalability [8] coded videos. These schemes use one- or multi-dimensionally *hierarchical scalability* provided by coding technologies, so that the user can obtain an image or a video at the permitted quality from one common codestream. *Hash chain* [9], [10] has also been introduced to several schemes for reduction of the number of managed keys and the keys delivered to the user [3]–[6].

Some hash chain-based access control schemes have been proposed for multimedia content. One of the schemes [11] manages a single key and also delivers a single key to each user. Although the conventional scheme controls access to each medium in multimedia content, only a single medium has a scalable hierarchy. Another scheme [12] can control access to not only media but also scalable hierarchies in each medium. In this scheme, however, the number of managed keys and delivered keys increases, depending on the number of media in multimedia content.

In this paper, we propose an efficient key derivation scheme for multimedia access control. The proposed scheme assumes that multimedia content consists of several media and there is a scalable hierarchy in each medium. By introducing a recursive hash chain, the number of managed keys is reduced to one and the number of delivered keys is also less than the conventional scheme [12]. The single managed key is not delivered to any user, providing security against key leakage. The proposed scheme is also resilient to collusion attacks, in which malicious users illegally access the multimedia content at higher quality than that allowed by their access rights.

This paper is organized as follows. Section II mentions the conventional access control schemes for multimedia content [11], [12] and describes the problems of the conventional schemes. The new scheme is proposed in Section III, and is analyzed in Section IV. Finally, conclusions are drawn in Section V.

II. Access Control for Multimedia Content

This section briefly describes the conventional access control schemes for multimedia content [11], [12], and summarizes the requirements for access control to clarify the aim of this work.

A. Conventional Schemes

1) Access Control for A Single Scalable Hierarchy [11]: The conventional scheme [11] assumes that multimedia content consists of M different media (image, video, audio, text, and so on) and that there is a scalable hierarchy (image/video resolution, frame rate, audio quality, etc) in a single medium. The scheme uses a symmetric encryption technique.

This scheme manages a single key. The first medium, where $m = 1$, has a scalable hierarchy. Encryption keys for entities in the first medium are derived from managed key K_1^1. Encryption keys K_1^d's are derived through a hash chain as

$$K_1^d = H^{d-1}\left(K_1^1\right), \ d = 2, 3, \ldots, D, \tag{1}$$

where $H^{\alpha}(\beta)$ represents a cryptographic one-way hash function $H(\cdot)$ applied to β recursively α times, and D represents the number of entities in the medium, i.e., the depth of the hierarchy. The d-th entity in the first medium is encrypted with its corresponding encryption key, K_1^d.

For the m-th medium where $m = 2, 3, \ldots, M$, keys K_m's are derived from key K_1^d through a recursive hash chain as

$$\begin{aligned} K_m &= II\left(f\left(K_1^d, II\left(K_1^d\right)\right)\right) \\ &= H\left(f\left(K_1^d, K_1^{d+1}\right)\right), \end{aligned} \tag{2}$$

respectively, where $f(\cdot)$ is a function with two inputs and one output in which the length of inputs and output are identical. A bitwise exclusive or (XOR) operation is a simple example of function $f(\cdot)$. As shown in Eq. (2) which represents a recursive hash chain introduced in this scheme, keys given previously are repeatedly used to derive another hash chain that is different from the ordinary hash chain.

Each user receives only a single key, and also receives the common encrypted multimedia content. If a user receives key K_1^{Δ} for the first medium, the user derives keys for accessible entities through the same hash chains as used in the encryption key derivation. These are,

$$K_1^{\delta} = H^{\delta-\Delta}\left(K_m^{\Delta}\right), \ \delta = \Delta+1, \Delta+2, \ldots, D, \tag{3}$$

and,

$$K_m = H\left(f\left(K_1^{\delta}, K_1^{\delta+1}\right)\right). \tag{4}$$

By using the decryption keys, the user decrypts entities which permitted to access. On the other hand, if a user receives key K_m for the m-th medium where $m = 2, 3, \ldots, M$, the user cannot derive any keys from K_m and can decrypt only the entity corresponding to key K_m.

2) Access Control for Multiple-Scalable Hierarchy [12]: Another conventional scheme [12] assumes that multimedia content consists of M different media (image, video, audio, text, and so on), in which a scalable hierarchy (image/video resolution, frame rate, audio quality, etc) exists; In the text medium, the appearing order of paragraphs has its own meaning, and it is referred to as a *semantic* hierarchy. This scheme also uses a symmetric encryption technique.

For a particular multimedia content consisting of M different media, this scheme manages M keys. Figure 1 shows an example of multimedia content where $M = 2$. For the m-th medium where $m = 1, 2, \ldots, M$, all encryption keys are derived from managed key K_m^1. Encryption keys $K_m^{d_m}$'s are derived through a hash chain as

$$K_m^{d_m} = H^{d_m-1}\left(K_m^1\right), \ d_m = 2, 3, \ldots, D_m+1, \tag{5}$$

where D_m represents the number of entities in the medium, i.e., the depth of the scalable hierarchy. The d_m-th entity in the m-th medium is encrypted with its corresponding encryption key, $K_m^{d_m}$.

Each user receives different set of M decryption keys due to which media/entities the user is allowed to access to, and also receives the common encrypted multimedia content. From the delivered keys, the user derives keys for accessible entities in accessible media through the same hash chain as used in the encryption key derivation. That is,

$$\begin{aligned} K_m^{\delta_m} &= H^{\delta_m-\Delta_m}\left(K_m^{\Delta_m}\right), \\ \delta_m &= \Delta_m+1, \Delta_m+2, \ldots, D_m, \end{aligned} \tag{6}$$

where $K_m^{\Delta_m}$ is the delivered key for the m-th medium. By using Δ_m decryption keys, the user decrypts Δ_m entities from the first entity to the Δ_m-th entity.

A user who receives $K_m^{D_m+1}$ cannot access any entities in the m-th medium, because one-way property of $H(\cdot)$ prevents the user to derive any other valid keys for the m-th medium of the multimedia content. The conventional scheme introduced this *unusable key* concept in order to cope with medium-based access control.

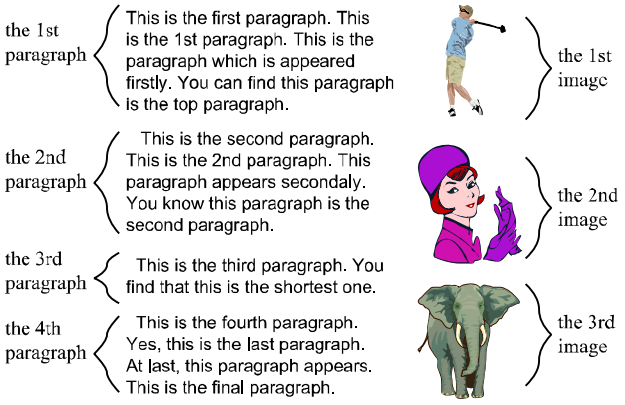


Fig. 1. An example of multimedia content (the number of media $M = 2$, the number of entities in the first medium $D_1 = 4$, and the number of entities in the second medium $D_2 = 3$).

B. Requirements

We describe four requirements for access control of multimedia content, i.e.,

- flexible access control,
- less number of managed and delivered keys,
- protection of managed key,
- collusion attack-resilience.

As mentioned in the previous section, the conventional scheme [11] can control access to only a single scalable hierarchy. It should be able to control access to multiple scalable hierarchies.

Although the conventional scheme [12] satisfy the first requirement, it encrypts entities in a medium independently of those in other media. This feature of the conventional scheme requires the same number of managed and delivered keys as the number of media in the multimedia content, i.e., M keys are managed and M keys are delivered to a user for the multimedia content consisting of M different media. This conventional scheme employs a simple hash chain [9] rather than cross-way hash trees [10].

The conventional schemes [11], [12] deliver the managed keys to users who are allowed to access some media at the

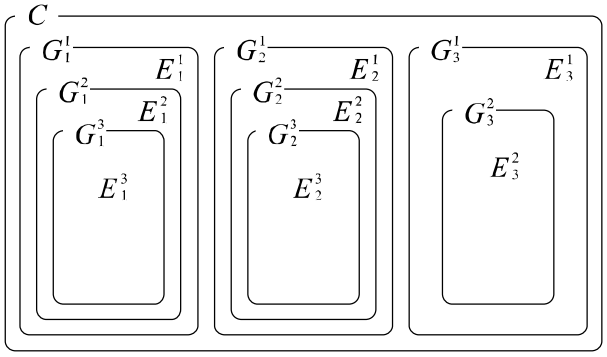


Fig. 2. An example of multimedia content conceptual diagram with multiple hierarchies (the number of media $M = 3$ and the depths in each medium $D_1 = 4$, $D_2 = 3$, and $D_3 = 2$).

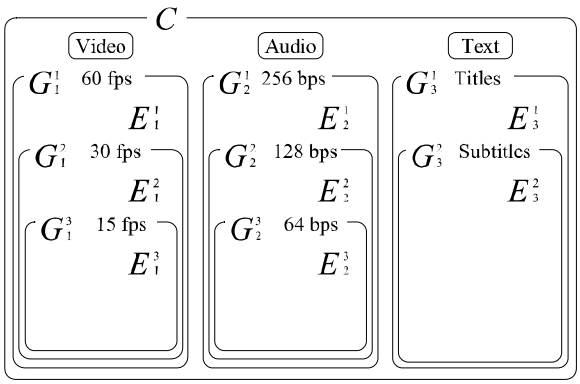


Fig. 3. A practical example of multimedia content with multiple hierarchies (the number of media $M = 3$ and the depths in each medium $D_1 = 3$, $D_2 = 3$, and $D_3 = 2$).

highest quality. The managed keys should not be delivered to any users and should be protected against key leakage.

A collusion attack is made by multiple users to obtain multimedia content with higher quality than that allowed by their access rights. For example, when a user who is allowed to display images and another user who is allowed to read texts share their keys, they can also obtain audio coupled with images and text paragraphs. Access control schemes must be resilient to collusion attacks.

In the next section, we propose a new key derivation scheme for multimedia access control. The proposed scheme assumes that there is a scalable hierarchy in each medium. This scheme manages only one key for a particular multimedia content and delivers less key to each user than the conventional scheme [12], regardless of which media/entities in the multimedia content the user can access. The single managed key is not delivered to any user. The proposed scheme is also resistant to collusion attacks.

III. Proposed scheme

First, we assume that multimedia content C consists of M media and each medium has a hierarchical structure;

$$C = \left\{G_1^1, G_2^1, \ldots, G_m^1, \ldots, G_M^1\right\}, \tag{7}$$

$$G_m^1 \supset G_m^2 \supset G_m^3 \supset \cdots \supset G_m^{D_m}, \quad m = 1, 2, \ldots, M, \tag{8}$$

where G_m^1 represents the m-th medium itself, and D_m is the depth of the hierarchy in the m-th medium. The complementary sets represent entities in medium G_m^1 as

$$E_m^{d_m} = G_m^{d_m} - G_m^{d_m+1}, \quad d_m = 1, 2, \ldots, D_m - 1, \tag{9}$$

and

$$E_m^{D_m} = G_m^{D_m}. \tag{10}$$

The proposed scheme derives keys from single managed key K_C and encrypts multimedia content C by encrypting $E_m^{d_m}$'s using those corresponding keys.

Fig. 2 shows an example conceptual diagram of the assumed multimedia content, where multimedia content C consists of three media, G_1^1, G_2^1, and G_3^1, i.e., $M = 3$, and the depths of each medium G_m^1 are three and two ($D_1 = 3$, $D_2 = 3$, and $D_3 = 2$), respectively, i.e.,

$$G_1^1 \supset G_1^2 \supset G_1^3, \tag{11}$$

$$G_2^1 \supset G_2^2 \supset G_2^3, \tag{12}$$

$$G_3^1 \supset G_3^2. \tag{13}$$

$E_m^{d_m}$'s are entities in medium G_m^1.

For easy understanding, more practical example of Fig. 2 is given in Fig. 3. Multimedia content C in Fig. 3 consists of video G_1^1, audio G_2^1, and text G_3^1, i.e., $M = 3$, and each medium has hierarchy whose depths are three, and two, i.e., $D_1 = 3$, $D_2 = 3$, and $D_3 = 2$, respectively.

A. Key Derivation and Encryption

In the example based on Fig. 3, access control is provided based not only on media, but also on each hierarchy in each medium. Keys for encryption are derived as shown in Fig. 4, and each key is used to encrypt and decrypt the corresponding entity. For example, $K_{E_1^1}$ is a key for entity E_1^1 which represents video frames decoded only at 60 fps. $K_{E_1^2}$ and $K_{E_1^3}$ are keys for E_1^2 and E_1^3, respectively. Keys $K_{E_2^{d_2}}$ and $K_{E_3^{d_3}}$ are similarly for the audio $E_1^{d_2}$ and the text $E_1^{d_3}$ ($d_2 = 1, 2, 3$, $d_3 = 1, 2$), respectively. It is noted that key K_C is the single managed key.

Firstly, key $K_{E_1^0}$ is derived from K_C as

$$K_{E_1^0} = H\left(K_C\right), \tag{14}$$

where $H(\cdot)$ is a cryptographic one-way hash function. Similarly, keys $K_{E_m^{d_m}}$'s are derived by

$$K_{E_m^{d_m}} = H^{d_m}\left(K_{E_m^0}\right), \quad d_m = 1, 2, \ldots, D_m, \quad m = 1, 2, 3, \tag{15}$$

where keys $K_{E_2^0}$ and $K_{E_3^0}$ are given in the next paragraph. Eq. (15) represents an ordinary hash chain [9], and the chain is shown with solid arrows in Fig. 4.

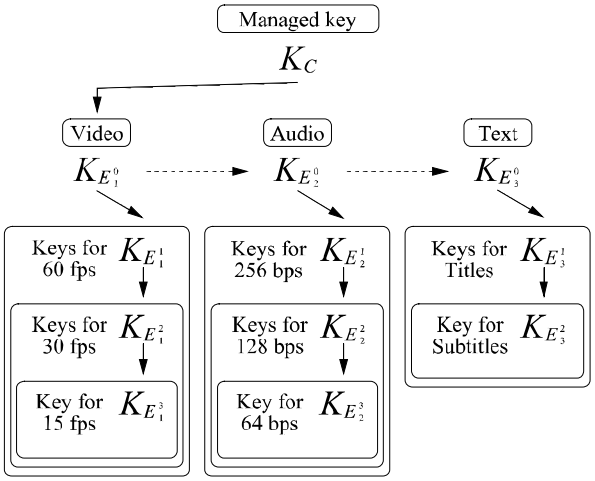


Fig. 4. Key derivation to control access to the multimedia content shown in Fig. 3. Solid arrows represent ordinary hash chains and dashed arrows represent a recursive hash chain.

Meanwhile, keys $K_{E_2^0}$ and $K_{E_3^0}$ are derived by a recursive hash chain. In this example, these keys are given as

$$\begin{aligned} K_{E_m^0} &= H\left(f\left(K_{E_{m-1}^0}, H\left(K_{E_{m-1}^0}\right)\right)\right) \\ &= H\left(f\left(K_{E_{m-1}^0}, K_{E_{m-1}^1}\right)\right), \\ m &= 2, 3, \end{aligned} \tag{16}$$

respectively, where $f(\cdot)$ is a function with two inputs and one output in which the length of inputs and output are identical. A bitwise exclusive or (XOR) operation is a simple example of function $f(\cdot)$. As shown in Eq. (16) which represents a recursive hash chain introduced in this scheme, keys given previously are repeatedly used to derive another hash chain that is different from the ordinary hash chain. The recursive hash chain is shown with dashed arrows in Fig. 4.

Each entity $E_m^{d_m}$ is encrypted using each corresponding key $K_{E_m^{d_m}}$, and then, multimedia content C is opened to public.

B. Delivered Keys and Decryption

1) User allowed to access three media: A user allowed to access the whole multimedia content receives three keys $K_{E_1^1}$, $K_{E_2^1}$, and $K_{E_3^1}$ as shown in Fig. 5 (a). The user derives all keys needed to decrypt all entities, through ordinary hash chains. Each user allowed to access three media at arbitrary quality also receives three keys $K_{E_1^{d_1}}$, $K_{E_2^{d_2}}$, and $K_{E_3^{d_3}}$. We assume that a user allowed to access each medium at the lowest quality, i.e., video at 15 fps, audio at 64 bps, and subtitle data. The user receives three keys $K_{E_1^3}$, $K_{E_2^3}$, and $K_{E_3^2}$ as shown in Fig. 5 (b). The user cannot, however, derive any keys from his/her delivered keys.

2) User allowed to access two media: Fig. 5 (c) shows an example user allowed to access two of the three media. In this example, the user can access video at 30 fps and title data. The user receives two keys $K_{E_1^2}$ and $K_{E_3^1}$, and derives keys $K_{E_1^3}$ for video and $K_{E_3^2}$ for text, respectively.

3) User allowed to access a single medium: If a user can access only movie at 30 fps, the user receives single key $K_{E_1^2}$ and derives key $K_{E_1^3}$ dependently as shown in Fig. 5 (d). Each user who can access a single medium receives single key $K_{E_1^{d_1}}$, $K_{E_2^{d_2}}$, or $K_{E_3^{d_3}}$.

In this scheme, the number of keys which a user receives is equal to the number of media which he/she can decode. Each user uses only ordinary hash chains to derive keys from the delivered keys. Keys K_C, $K_{E_1^0}$, $K_{E_2^0}$, and $K_{E_3^0}$ are not delivered to any user.

C. Features

Four main features of the access control scheme are briefly summarized here. They have satisfied with the requirements described in Section II-B.

The proposed scheme realizes flexible access control for multimedia content. It can control access to not only media but also scalable hierarchies in each medium.

This scheme, introducing a recursive hash chain, has reduced the number of managed keys to one, which is as many as the conventional scheme [11]. The number of delivered keys also is less than the conventional scheme [12] which manages and delivers the same number of keys as media in the multimedia content.

Each key for each entity is derived from the single managed key. The managed key is not delivered to any user.

The proposed scheme using a recursive hash chain can prevent malicious users to collude to decode multimedia content at higher quality than that allowed by their access rights.

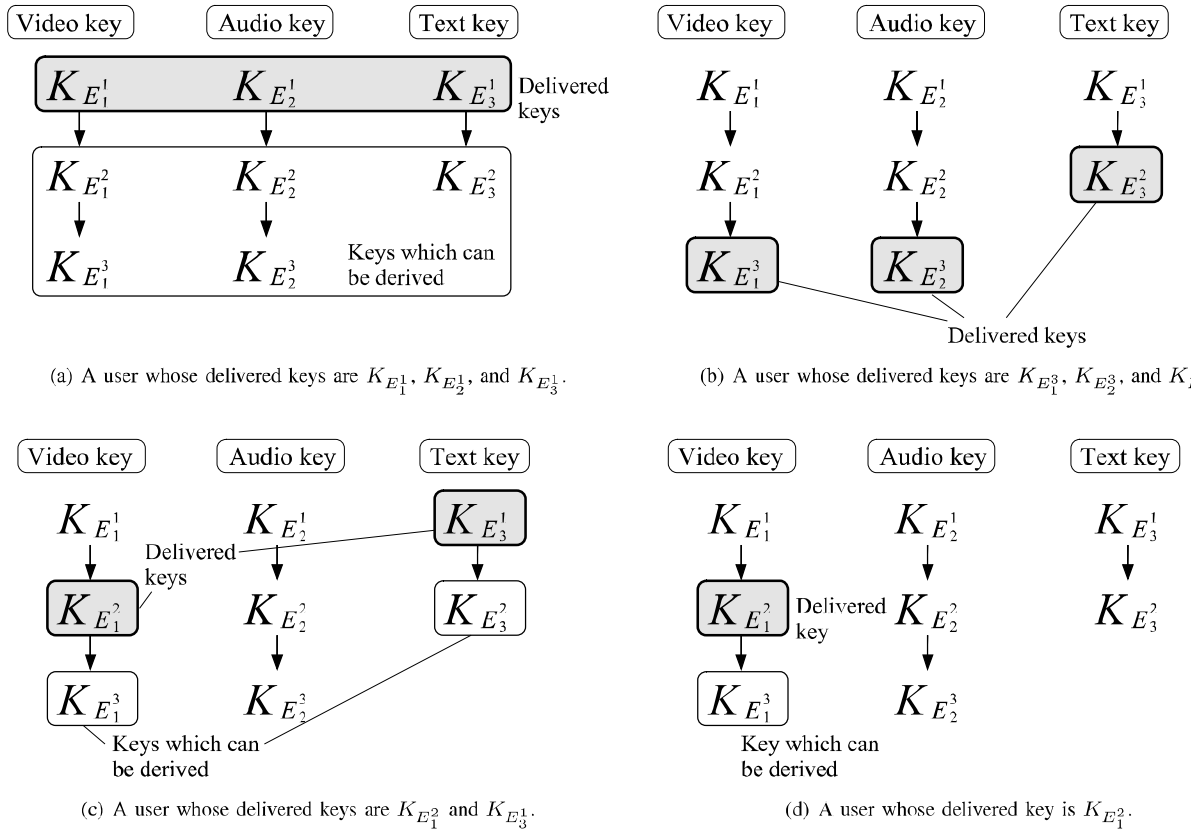
It is noted that any arbitrary function and key combination can be used for a recursive hash chain. In addition, it is noted that any arbitrary key assignment can be used to properly control access to the multimedia content.

IV. Evaluation

The proposed scheme using a recursive hash chain is evaluated by comparing with the conventional schemes [11], [12]. Evaluation is given in terms of the number of media with a scalable hierarchy, the number of managed and delivered keys, protection of managed keys, and collusion attack resilience.

Table I shows the results of comparisons. The proposed scheme and the conventional scheme [12] assume that there is a scalable hierarchy in each medium, whilst the conventional scheme [11] assumes that there is a scalable hierarchy in only a single medium.

The proposed scheme and the conventional scheme [11] manages only a single key regardless of the number of media and the depths in each or one medium, whilst the conventional scheme [12] must manage the same number of keys as media in the multimedia content. The proposed scheme also delivers the same number of keys as accessible media, while the conventional scheme [12] should deliver the same number of keys as media in the content. The conventional scheme [11] delivers a single key to each user, because this scheme assumes that only a single medium has a hierarchical structure.



(a) A user whose delivered keys are $K_{E_1^1}$, $K_{E_2^1}$, and $K_{E_3^1}$.

(b) A user whose delivered keys are $K_{E_1^3}$, $K_{E_2^3}$, and $K_{E_3^2}$.

(c) A user whose delivered keys are $K_{E_1^2}$ and $K_{E_3^1}$.

(d) A user whose delivered key is $K_{E_1^2}$.

Fig. 5. Delivered keys and derived keys for each user.

TABLE I
COMPARISONS IN TERMS OF (I) THE NUMBER OF MEDIA WITH A SCALABLE HIERARCHY, (II) THE NUMBER OF MANAGED KEYS, (III) THE NUMBER OF DELIVERED KEYS, (IV) PROTECTION OF MANAGED KEYS, AND (V) COLLUSION ATTACK RESILIENCE.

	Proposed	Conventional [11]	Conventional [12]
I	M	1	M
II	1	1	M
III	between 1 and M	1	M
IV	Yes	No	No
V	Yes	Yes	Yes

The single managed key is not delivered to any user in the proposed scheme, whereas the managed keys are delivered to some users in the conventional schemes [11], [12]. The proposed scheme is also as resilient to collusion attacks as the conventional schemes [11], [12]. The table brings out the effectiveness of the proposed scheme.

V. CONCLUSION

This paper has proposed an efficient key derivation scheme for multimedia access control, in which a recursive hash chain is employed. The proposed scheme can control access to a scalable hierarchy not only in a single medium, but also in each medium, and achieves more flexible access control than the conventional scheme. The scheme manages only a single key regardless of the number of media and the depths in each medium. Each user also receives less keys than the conventional scheme. The single managed key is not delivered to any user. This proposed scheme also prevents collusion attacks, in which malicious users illegally access the multimedia content at higher quality than that allowed by their access rights.

REFERENCES

[1] B. B. Zhu, M. D. Swanson, and S. Li, "Encryption and authentication for scalable multimedia: current state of the art and challenges," in *Proc. SPIE*, vol.5601, pp.157–170, 2004.
[2] Z. Shahid, M. Chaumont, and W. Puech, "Selective and scalable encryption of enhancement layers for dyadic scalable H.264/AVC by scrambling of scan patterns," in *Proc. IEEE ICIP*, pp.1273–1276, 2009.
[3] Y. Wu, D. Ma, and R.H. Deng, "Progressive protection of JPEG 2000 codestreams," in *Proc. IEEE ICIP*, pp.3447–3450, 2004.
[4] Y. G. Won, T. M. Bae. and Y. M. Ro, "Scalable protection and access control in full scalable video coding," in *Proc. IEEE IWDW*, pp.407–421, 2006.
[5] S. Imaizumi, M. Fujiyoshi, Y. Abe, and H. Kiya, "Collusion attack-resilient hierarchical encryption of JPEG 2000 codestreams with scalable access control," in *Proc. IEEE ICIP*, pp.II–137–II–140, 2007.
[6] S. Imaizumi, M. Fujiyoshi, and H. Kiya, "Efficient collusion attack-free access control for multidimensionally hierarchical scalability content," in *Proc. IEEE ISCAS*, pp.505–508, 2009.
[7] *Information technology — JPEG 2000 image coding system – Part 1: Core coding system.* ISO/IEC 15444–1, 2004.
[8] *Information technology — Coding of audio – Visual objects – Part 2: Visual.* ISO/IEC 14496–2, 2004.

[9] L. Lamport, "Password authentication with insecure communication," *Communications of the ACM*, vol.24, no.11, pp.770–772, 1981.
[10] M. Joye and S. M. Yen, "One-way cross-trees and their applications," in *Proc. IACR PKC*, pp.355–358, 2002.
[11] S. Imaizumi, M. Fujiyoshi, and H. Kiya, "An efficient access control method for composite multimedia content," *IEICE Electronics Express*, vol.7, no.20, pp.1534–1538, 2010.
[12] M. Fujiyoshi, W. Saitou, O. Watanabe, and H. Kiya, "Hierarchical encryption of multimedia contents for access control," in *Proc. IEEE ICIP*, pp.1977–1980, 2006.