INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 MPEG2016/m39152
October 2016, Chengdu, China

Source **Thales, IETR-INSA, Telecom ParisTech**
Status **Input Document**
Title **Additional box in ISOBMFF for selective encryption**
Author Cyril Bergeron, Wassim Hamidouche, Jean Le feuvre, Cyril Concolato

1 Introduction

This contribution aims to bring use cases of Visual identity management (also described in details in [1]). As shown in Figure 1, in this our case, media streaming servers are offered conditional access for providing streaming which is partially encrypted. Users who join the real-time streaming session, they can decrypt some people by using privacy key. But the actual video format and signalization do not manage partial encryption (i.e. when the video or regions of the video is encrypted the whole video can not be decoded). However, in such uses-cases, the user may be interested by parts of the video even when other parts are scrambled. In this contribution we propose a new scheme protection scheme box enabling such new features.

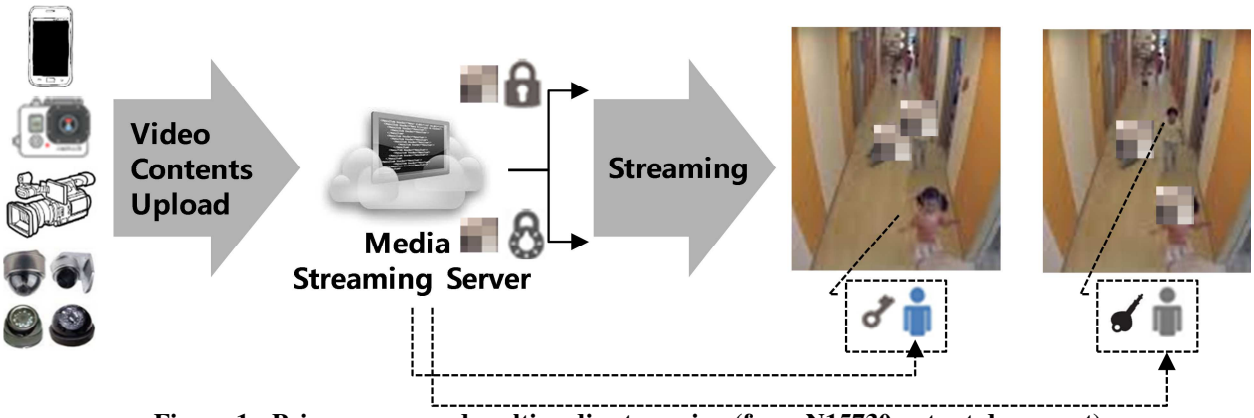

Figure 1 - Privacy managed multimedia streaming (from N15730 output document)

2 Reminder about selective encryption

The most straightforward method to secure video content is to encrypt the whole bitstream using standard encryption algorithms such as Advanced Encryption Standard (AES). This method called Naive Encryption Algorithm (NEA) treats the video bitstream as text data without

considering the structure of the compressed video. However, NEA suffers from several drawbacks:

1. The encryption/decryption process becomes time and energy consuming (computationally costly) for large-scale data, especially video at high resolution and high bitrate, in particularly for mobile devices.

2. The NEA prevents untrusted middle-box in the network to perform post-processing operations on the encrypted video bitstream such as repackaging and transmuxing.

3. The NEA solution applied on the scalable video bitstream does not preserve the bitstream features such as sub-stream extraction for network adaptation and error resilience. This also may cause an excepted behaviour of the video decoder since the encrypted video bitstream is not standard format compliant.

Selective video encryption has emerged as an effective alternative to NEA. Selective video encryption considers the coding structure of the video bitstream and encrypts only the most sensitive information in the video bitstream. Selective encryption solutions can guarantee the five following requirements:

1- The encrypted bitstream remains compliant with the standard (can be decoded by standard decoders).

2- The encryption algorithm does not affect the compression ratio.

3- Achieve a high security level with a minimum additional delay and computational complexity.

4- Maintain the video bitstream features including secure mid-network adaptation and error resilience.

5- Provide different levels of security: visual encryption, perceptual encryption and ROI encryption.

The key issue is how to select the sensitive data to encrypt. During video encoding, sensitive data like value of symbols of intra-prediction mode, residual data and motion vector difference could be partially encrypted. It provides an approach for selecting sensitive data to encrypt which makes it time efficient, secure and format compliant. Without other constraints, the drawback of this simple method is that the selective encryption brings bitrate increase since encryption "randomizes" coded symbols and reduces the entropic coding efficiency.

In this contribution, we focus on a Selective Encryption scheme which operates in compressed domain particularly at the entropic (de-)coder level. Thus, it overcomes the previous drawback: it does not affect the bitrates and has the advantage of being suitable for streaming over heterogeneous network because of no change in bitrates and in video format compliance. Indeed, Selective Encryption does not pursue message privacy protection. Instead, the method modifies the message content in such a way that an unauthorized receiver (i.e. a standard decoder) can normally decode any ciphered stream, while the displayed content is made non intelligible by the

encryption. Selective encryption methods aim to cipher video data without disrupting a standard decoding process.

3 Selective encryption for region encryption

ISO/IEC 23001-7 (Common encryption in ISO base media file format files) and ISO/IEC 14496-12 (ISO base media file format) define different modes and signaling of encryption for compressed stream. The ISO/IEC 23001-7 enables signaling to encrypt a part of the picture (region), however the access unit is indicated as not decodable without correct decryption (the video can not be decoded since the encrypted bitstream is not format compliant).

Since the selective encryption takes place inside CODEC, it is also possible to restrain the selective encryption locally (i.e. on selected Macroblock or block) and/or temporally (i.e. frame) by the segmentation of frames in slices, and by the reference picture management to prevent propagation of visual scrambling on unwanted portion of images (Figure 2).

Figure 2 - spatial ciphering

For this use case, we can use in AVC ‘slice_groups’ (as described in Annex A in subclause G.10 in ISO/IEC 14496 10), which defines a subset of the macroblocks in a coded picture and may contain one or more slices. It also permits to map the sequence of coded MBs to the decoded picture in a number of flexible ways. The allocation of macroblocks is determined by a slice_group_map_type that indicates which slice_group each MB belongs to, in Picture Parameter Set data (see subclause 7.4.2 in ISO/IEC 14496-10). An example of 2 slice_groups is shown in Figure 3, with map_types 2 (also called ‘Foreground and Background’).

Figure 3 -Frame with 2 slice groups

Here, we can considerate that the foreground (slice_group1) will be ciphered by selective encryption and the background (slice_group0) will be unencrypted. In the case of the HEVC standard, the ROI encryption can be performed using tile encryption with restriction of Motions Vectors inside the encrypted parts to prevent the propagation of encryption outside the encrypted region [2].

4 Support for Protected Streams in ISOBMFF

In ISOBMFF, protected content is supposed to be a file-format transformation. But when the content has been transformed (e.g. by encryption) in such a way, it can no longer be decoded by the standard decoder. The transformation functions by encapsulating the original media declarations. The encapsulation changes the four character-code of the sample entries, so that protection-unaware readers see the media stream as a new stream format.

However, in the case of selective encryption, “protected” streams is only visually scrambled while keeping it fully decode-able even by a non-robust standard compliant decoder. So the use of Protection Scheme Information Box (‘sinf’) is not adapted to our constraints.

5 Proposal: Scramble Scheme Information Box

The proposition of syntax is similar to ‘sinf’ since it reuses Scheme Type Box (‘schm’) and Scheme Information Box (‘schi’) to integrate Common Encryption signalizations.

Definition
Box Types: 'scrb''
Container: Protected Sample Entry, or ItemProtectionBox
Mandatory: Yes
Quantity: One or More

The ScrambleSchemeInfoBox contains all the information required both to understand the scramble operation applied and its parameters, and also to find other information such as the kind and location of the key management system. The ScrambleSchemeInfoBox is a container Box. It is mandatory in a sample entry that uses a code indicating a protected stream for visually scrambling.

Syntax:

```
aligned(8) class ScrambleSchemeInfoBox(fmt) extends Box('scrb')
{
SchemeTypeBox scheme_type_box;
SchemeInformationBox info; // optional
}
```

6 Conclusions

This contribution presented a way of signaling visual scrambling to address its use case. We recommend that MPEG experts take into consideration this solution for adding it in the scope of ISOBMFF

7 References

[1] w16226 - Use cases and requirements on Visual Identity Management AF (June 2016, Geneva,CH).

[2] Mousa Farajallah, Wassim Hamidouche, Olivier Déforges, Safwan El Assad, ROI encryption for the HEVC coded video contents, IEEE conference on Image Processing ICIP 2015.