A Survey of H.264 AVC/SVC Encryption-Part III

IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 22, NO. 3, MARCH 2012

V. Recommended Solutions

- *A. Highest Level Security*
- Raw video data must be encrypted before applying any other kinds of processing , such as compression, on it.
- The main concern is:
- Can it still be decryptable if there is a lossy compression involved in between the encryption and the decryption modules?

- The second issue is:
- Even if a lossless compression is applied, the compressibility of the encrypted data is questionable because the correlation among raw data will be distorted by the encryption process. Therefore, only a very small part of the raw data can be encrypted before compression is applied.

- If compression before encryption is allowed: Solutions can target the encoding process, i.e., compress the raw data to a certain source-independent bitrate and encryption is applied directly on the compressed bit-stream.

- A practical solution for SVC defines target bitrates for the base and enhancement layers, and packetizes base and enhancement layers into fixed length packets and encrypts them independently.

- The proposed solution is as secure as the employed encryption primitive, e.g., AES in counter mode, with respect to Message Preserving security on the raw video data.

- Since Encryption is applied on the compressed bitstream, the runtime performance is as good as the runtime performance of the applied encryption primitive, which is very good in the case of AES.

- Format-compliance can not be preserved, optimal packetization leads to a violation of the MP security notion on raw video data. Thus there is a tradeoff between frame structure or scalability preservation and compression complexity and efficiency.

- Transcoding and DCT watermarking can not be conducted directly in the encrypted domain, but stream substitution watermarking is still possible. For stream substitution watermarking we propose the application of a stream cipher mode with no cipher feedback.
- The substitution watermark in the encrypted domain, w_e , is then $w_e = w_p \oplus k$, where w_p is the plaintext watermark and k are the corresponding key stream bits.

- *B. Content Confidentiality (A specific requirement of multimedia security)*
- To securely implement content confidentiality it is necessary to encrypt most of the video data or to prevent (partial) decoding at all.

- For this application scenario we propose to encrypt the NALU payload, because each of the compression-integrated encryption schemes leaks some visual information. A combination of compression-integrated schemes heavily distorts the videos (see Fig. 8) but a human observer can guess the content of the video--- approximation attack.

- The encryption algorithm can be designed to preserve the NALU syntax and semantics, so as to preserve format-compliance or certain functionalities.

- Another approach is the application of *Container* formats, which more clearly separate encryption and compression, but introduce certain overheads in terms of runtime and compression efficiency (e.g., results for adaptation for a NAL-based system and a MPEG-21-based system are presented [47], that can serve as reference for the expected overhead if *Container* formats are employed).

- [47] R. Kuschnig, I. Kofler, M. Ransburg, and H. Hellwagner, “Design options and comparison of in-network H.264/SVC adaptation,” *J. Vis. Commun. Image Representation*, vol. 19, no. 8, pp. 529–542, Dec. 2008.

- Partial encryption schemes that encrypt a subset of NALUs are not applicable for content confidentiality as edge images can be reconstructed [see Fig. 8(e)–(g)]. Partial encryption schemes that encrypt the start of NALUs (i.e., header files) may offer some security with respect to content confidentiality.

- In case of CAVLC only the coefficient data is coded context-adaptively, and thus at least some coded coefficient data has to be encrypted. For CAVLC the security relies on the uncertainty of the number of non-zero coefficients in neighboring blocks, at most 17 × 17 combinations (0 to 16 non-zero coefficients) have to be considered for single 4 × 4 DCT transformed residual block, not an easy, but solvable task, using back tracking algorithms that try to correctly decode the partial coefficient data.

- For CABAC the security relies on the uncertainty of the state of arithmetic decoding, which means that the current state in decoding (which syntax elements to decode) and the state of each of the approximately 400 contexts have to be guessed, as well as the state of the arithmetic decoding variables codIRange and codIOffset (both represented with 16 bit).

- Overall a very complex problem an attacker has to solve in order to obtain a partial decoding of the slice data. Thus, although there is no formal proof of the security of these schemes with respect to content confidentiality, a potential adversary is assumed to have a hard time to decode partially encrypted CABAC encoded video data.

- *C. Sufficient Encryption*
- Most proposed schemes are suitable for sufficient encryption of H.264, the relaxed security requirements make it possible to employ encryption schemes, that offer functionalities that would violate the security requirements of other security and application scenarios. Contrary to transparent encryption there is no minimal quality requirement for the cipher video, which makes sufficient encryption easier to implement.

- In case of SVC we propose the encryption of the base layer [see Fig. 8(f) and (g) for replacement attacks, i.e., we have replaced the encrypted picture data by pictures containing only zero values]. If format-compliance is desired, the base layer can be replaced by a uniformly grey video sequence (negligible compression performance deficits).

- In the case of CAVLC partial decoding is a more realistic threat then in the case of CABAC and thus the application of CABAC and the encryption of the leading fraction of the NALUs including several bits (>>128) of the arithmetically coded data is proposed. In the case of CABAC the remaining fraction is hard to decode as all the internal states of the CABAC engine have to be known.

- Security proofs for sufficient encryption are not found in literature, but successful attacks against previously proposed schemes have been reported. Due to the application of unreliable quality metrics for low quality visual data, e.g., peak signal-to-noise ratio (PSNR) does not perform well for that purpose [48], [49], the results of the attacks remain rather incomparable.

- [48] T. St¨utz, V. Pankajakshan, F. Autrusseau, A. Uhl, and H. Hofbauer, “Subjective and objective quality assessment of transparently encrypted JPEG2000 images,” in *Proc. ACM MMSEC*, Sep. 2010.
- [49] H. Hofbauer and A. Uhl, “Visual quality indices and low quality images,” in *Proc. IEEE 2nd Eur. Workshop Visual Inform. Process.*, Jul. 2010, pp. 171–176.

- Compression-integrated encryption often is associated with a severe reduction of compression performance, however, many of the proposed schemes for H.264 perform very well (see Table II for an overview).

- *D. Transparent Encryption:* The main additional requirement of transparent encryption is quality control.
- Though many schemes have been proposed under the label perceptual encryption, quality control of the encrypted data is only discussed in a few contributions [50].
- [50] E. Magli, M. Grangetto, and G. Olmo, "Conditional access techniques for H.264/AVC and H.264/SVC compressed video," *IEEE Trans. Circuits Syst. Video Technol.*, 2008.

- In case of AVC, a transparent encryption approach has been proposed in [51], employing restricted MVD encryption and the encryption of less important bitplanes of DCT coefficients.

- [51] E. Magli, M. Grangetto, and G. Olmo, “Conditional access to H.264/AVC video with drift control,” in *Proc. IEEE ICME*, Jul. 2006, pp. 1353–1356.

- Furthermore previous DCT sign and coefficient encryption proposals can be extended by an explicit quality control, which controls the quality by restricting the sign (CAVLC and CABAC) and magnitude (CABAC) encryption to certain coefficients and additionally only the magnitude could be encrypted.

- In [52], it is proposed to employ SVC for transparent encryption and if format-compliance is targeted to force a decoder to ignore the encrypted data by signalling in the NAL (unspecified NUTs). If the application of SVC is possible the encryption of the enhancement layers is the recommended solution [see Fig. 1(b) for an illustration of the approach].
- [52] T. St¨utz and A. Uhl, “Format-compliant encryption of H.264/AVC and SVC,” in *Proc. 8th IEEE ISM*, Dec. 2008.

- Security of transparent encryption schemes relies on the inability of an adversary to compute higher quality versions than already made public. Thus specifically tailored algorithms, inspired by super-resolution and denoising algorithms, are the main threat.

- In [51], only slight bitrate increases of less then 1% are reported for their transparent AVC encryption scheme.
- There is no compression overhead for the SVC-based encryption scheme.

- In case of the SVC encryption approach the qualities of the substreams have to be determined in the SVC compression process, which highly depends on the desired scalability properties of the SVC bitstream and is more complex then AVC encoding.
- If an SVC bitstream is available the scheme is efficient.

- Commonly format compliance is considered a necessity for the transparent encryption scenario and thus has to preserved. The proposed DCT watermarking schemes can not be applied, stream substitution watermarking can still be applied.

- *E. ROI Encryption*
- ROI encryption has been primarily proposed for privacy preserving encryption schemes.
- 1) *Suitable Video Encryption Schemes:* In [53] and [54], sign encryption and permutations is proposed and reported to meet the security constraints [53].

- [53] F. Dufaux and T. Ebrahimi, “A framework for the validation of privacy protection solutions in video surveillance,” in *Proc. IEEE ICME*, Jul. 2010, pp. 66–71.
- [54] F. Dufaux and T. Ebrahimi, “H.264/AVC video scrambling for privacy protection,” in *Proc. IEEE ICIP*, Oct. 2008, pp. 1688–1691.
- [55] S. Li, C. Li, G. Chen, N. Bourbakis, and K. Lo, “A general quantitative cryptanalysis of permutation-only multimedia ciphers against plaintext attacks,” *Signal Process.: Image Commun.*, vol. 23, no. 3, pp. 212–223, Jan. 2008.

- 2) *Security:* According to [53], sign encryption and permutation prevents automatic face recognition and this is their proposed security metric for privacy preserving encryption. The goal of an adversary for this security notion is the development of a face recognition system that can identify faces even when encrypted. An adversary will try to combine attacks against the video encryption scheme and improved face recognition systems, e.g., permutations are known to be susceptible to known plaintext attacks [55].

- 3) *Compression:* Only small decreases in compression complexity are reported for [53], [54].
- 4) *Complexity:* As the privacy-threatening regions (faces) have to be detected, which is done on the raw video data and can be treated as an online scenario. Thus the impact of the overall system complexity is small.

- 5) *Preserved Functionality:* An important feature for ROI encryption is that the remaining video can be decoded in sufficient quality, such that privacy-preserving surveillance is possible. The recommended schemes have this property and are also format-compliant [54].

VI. Discussion and Analysis

- H.264 encryption schemes are capable to preserve diverse functionality, but naturally at some cost in terms of security, runtime performance and compression performance.
- Table II summarizes important aspects and properties of the diverse encryption algorithms.

- In order to put the complexities of encryption and compression into perspective, we give experimental results for H.264 decoding and AES encryption, which clearly indicate that AES encryption is almost negligible compared to H.264 compression.
- On a Intel(R) Core(TM)2 CPU 6700 at 2.66 GHz the ffmpeg implementation is able to decode HDTV content (1080p) in high profile with CABAC (BluRay) at about 27 f/s.

- This is roughly equivalent to a data throughput of about 40 Mbits, which is about 30% higher if the baseline profile is applied. The application of CAVLC instead of CABAC increases the throughput about 15%, which is partly due to an increased access frequency of CABAC [58].
- [58] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockerhammer, and T. Wedi, “Video coding with H.264/AVC: Tools, performance, and complexity,” *IEEE Circuits Syst. Mag.*, vol. 4, no. 1, pp. 7–28, Jan.–Apr. 2004.

- Compared to AES encryption that achieves a data throughput of about 1159 Mbits (with the openssl implementation).
- Even if we consider that only partial decoding would be sufficient for most in-compression encryption schemes the complexity of partial decoding and re-encoding remains huge compared to AES encryption.

- The bitstream-parsing and entropy decoding portion (CAVLC) of H.264 decoding has been experimentally found to be about 13% of the overall decoding complexity [c1]. The CABAC decoding complexity is roughly the double for the ffmpeg implementation as compared to the CAVLC decoding complexity.

- [c1] M. Horowitz, A. Joch, F. Kossentini, and A. Hallapuro, “H.264/AVC baseline profile decoder complexity analysis,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 704–716, Jul. 2003.
- [c2] K. Denolf, C. Blanch, G. Lafruit, and A. Bormans, “Initial memory complexity analysis of the AVC codec,” in *Proc. IEEE Workshop SIPS*, Oct. 2002, pp. 222–227.
- [c3] T.-C. Chen, S.-Y. Chien, Y.-W. Huang, C.-H. Tsai, C.-Y. Chen, T.- W. Chen, and L.-G. Chen, “Analysis and architecture design of an HDTV720p 30 frames/s H.264/AVC encoder,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 6, pp. 673–688, Jun. 2006.

- Thus partial H.264 decoding and reencoding is substantial compared to AES encryption. In-depth analyses of the computational complexity of H.264 can be found in [cc],[c1]–[c3].
- For H.264 encryption these results indicate that apart from an online-scenario in which compression is performed anyways, there have to be good reasons (required functionality preservation) to apply a compression-integrated encryption scheme.

- Partial encryption can only make sense in an offline-scenario, in which no compression is performed, as otherwise the potential performance gains are negligible compared to the overall system’s complexity.

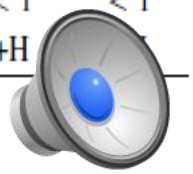
- Although a few proposals for quality/intelligibility/security metrics have been made [c4], [c5], none actually evaluates the performance of the proposed metrics with respect to the correlation to human perceived quality/intelligibility.

- [c4] Y. Mao and M. Wu, “A joint signal processing and cryptographic approach to multimedia encryption,” *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 2061–2075, Jul. 2006.
- [c5] J. Sun, Z. Xu, J. Liu, and Y. Yao, “An objective visual security assessment for cipher-images based on local entropy,” *Multimedia Tools Applicat.*, vol. 53, no. 1, pp. 75–95, Mar. 2010.

TABLE II

OVERVIEW OF H.264 ENCRYPTION APPROACHES

	Naive	MPV	CF	FC	NC	S	L	SDCT	MVD	SSO	Inter	Intra
Highest level security	X	✓	X	X	X	X	X	X	X	X	X	X
Content confidentiality	✓	✓	✓	✓	✓	X	X	X	X	X	X	X
Sufficient encryption	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Transparent encryption	X	X	X	SVC	SVC	~	~	~	~	~	~	~
FC transparent encryption	X	X	X	SVC	X	~	~	~	~	~	~	~
ROI encryption	X	X	X	X	✓	✓	✓	✓	X	✓	X	✓
Combinable	X	X	X	X	X	✓	✓	✓	✓	✓	✓	✓
Format-compliant	X	X	X	✓	X	✓	✓	✓	✓	✓	✓	✓
Compliant packetization	X	X	X	✓	✓	✓	✓	✓	✓	✓	✓	✓
Compliant adaptation	X	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adaptation	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Transcodability	X	X	X	X	X	✓	X	✓	✓	X	✓	X
DCT Watermarking	X	X	X	X	X	✓	X	✓	✓	~	✓	✓
BSS Watermarking	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Compression pres.	✓	~	~	✓	✓	~	✓	✓	~	~	✓	✓
Error propagation	H.264-Full	H.264	H.264	H.264	H.264	H.264	H.264	H.264	H.264	H.264	H.264	H.264
Complexity (online)	b/p	b/p + ϵ	b/p	b/p	b/p	< 1	$\approx b/p$	1	< 1	≈ 1	< 1	< 1
Add. Comp. (offline)	0	0	0	0	0	+H	+H	+H	+H	+H	+H	



VII. Outlook and Further Research Directions

- The most apparent deficit in the current research is that, although security in the context of video encryption is defined with respect to quality and intelligibility, neither quality nor intelligibility can be assessed.

- The lack of objective assessment methods makes video encryption schemes incomparable; the analysis on the basis of a visual inspection of single frames, as it is the current state-of-the-art can hardly be considered satisfactory in a scientific context.

- Thus further research should focus on the development of objective metrics for the assessment of the security of video encryption schemes for the different security and application scenarios. A prerequisite for the development of novel objective quality/security metrics are subjective tests, in which the actually perceived quality and intelligibility is determined by human observers.

- Contributions to this line of research have already been made for transparent JPEG2000 encryption [c6]. Objective assessment on the basis of face recognition rates has been proposed for the analysis of privacy preserving encryption [c7].

- [c6] T. St¨utz, V. Pankajakshan, F. Autrusseau, A. Uhl, and H. Hofbauer, “Subjective and objective quality assessment of transparently encrypted JPEG2000 images,” in *Proc. ACM MMSEC*, Sep. 2010.
- [c7] F. Dufaux and T. Ebrahimi, “A framework for the validation of privacy protection solutions in video surveillance,” in *Proc. IEEE ICME*, Jul. 2010, pp. 66–71.

- *A. Standardization Efforts*
- Further efforts in the area of H.264 encryption should also consider the standardization of security tools within H.264. Bitstream-oriented encryption, as well as other security features, such as authentication/message integrity and scalable authentication/message integrity, could be optimally integrated into the existing H.264 framework, as due to the well-designed NAL abstraction a backwards-compatible integration of security features would be possible.

- Such a transparent, backward-compatible integration of security tools (encryption/authentication/hashing), i.e., the secured bitstream is still H.264-compliant and its plaintext portions can be decoded by any H.264-compliant decoder, could turn out to be of ultimate importance and utility for practical application.

- A standard for security tools within H.264 should employ cryptographic primitives (symmetric ciphers, hash functions) as building blocks and leave the actual choice of algorithms (AES, SHA-256) open for further extensions (a good and common practice [c8]).

- Previous standardization efforts in the field of video encryption (MPEG-4 Part 13, MPEG-21 Part 4) have aimed to standardize generic frameworks for the implementation of digital rights management systems [c10]–[c12].

- [c9] *Advanced Video Coding for Generic Audiovisual Services*, ITU-T H.264, Nov. 2007.
- [c10] M. Ji, S. Shen, W. Zeng, T. Seno, and T. Ueno, “MPEG-4 IPMP extension: For interoperable protection of multimedia content,” *EURASIP J. Appl. Signal Process*, vol. 2004, no. 14, pp. 2201–2213, 2004.
- [c11] *Information Technology—Coding of Audio-Visual Objects, Part 13: Intellectual Property Management and Protection (IPMP) Extension*, ISO/IEC 14496-13, 2004.
- [c12] *Information Technology—Multimedia Framework (MPEG-21)—Part 7: Intellectual Property Management and Protection Components*, ISO/IEC 21000-4:2006, Apr. 2006.

- The security tools standardized within the scope of H.264 could be integrated into the existing MPEG-4 IPMP and MPEG-21 IPMP frameworks by the definition of the corresponding tools, i.e., standardization of H.264 security tools would complement, extend, and improve the applicability of the already published MPEG-4 IPMP and MPEG-21 IPMP standards.