

Selective Encryption of Multimedia Content in Distribution Networks: Challenges and New Directions

Xiliang Liu
The Graduate Center
The City University of New York
365 Fifth Avenue, New York, NY 10016, USA
Email: xliu@cs.cuny.cuny.edu

Ahmet M. Eskicioglu*
Department of Computer and Information Science
Brooklyn College of the City University of New York
2900 Bedford Avenue, Brooklyn, NY 11210, USA
Email: eskicioglu@sci.brooklyn.cuny.edu

ABSTRACT

The security of multimedia data in digital distribution networks is commonly provided by encryption, i.e., the mathematical process that transforms a plaintext message into unintelligible ciphertext. Nevertheless, the classical and modern ciphers have all been developed for the simplest form of multimedia data, i.e., text, and are not appropriate for higher forms such as images and video with very large file sizes. *Selective encryption* is a recent approach to reduce the computational requirements for huge volumes of multimedia data in distribution networks with different client device capabilities. In this paper, we provide a survey and classification of the proposed schemes, discuss the current issues and present some future directions.

Key words: Selective encryption, content protection, multimedia, cipher, JPEG, MPEG

1. INTRODUCTION

The widening Internet bandwidth and availability of digital consumer electronics (CE) devices for playback, recording and storage have increased the demand for multimedia services. In the digital domain, distribution networks need to address two fundamental problems: (1) Reduction of huge communication requirements for multimedia data, and (2) Protection of copyrighted multimedia data. A solution to the first problem is provided by efficient coding techniques for images, audio and video. Compression [JPEG, JPEG 2000, MPEG-1, MPEG-2, MPEG-4, H.26X] removes the spatial and temporal redundancy in multimedia data with imperceptible degradation. The second problem is addressed in privately defined closed systems by controlling access to copyrighted content [1]. The CE devices that receive satellite and cable transmissions are equipped with the software and hardware needed to prevent unauthorized access.

The Internet, however, remains to be a public network of networks with a vast potential for multimedia content distribution. Although the first examples of PC-based Digital Rights Management Systems (DRMs) are emerging in the commercial market [2], the real-time constraints for many multimedia applications cannot be met by an increasing number of smaller client devices (such as PDAs and videophones) with limited processing and communication power. Hence, there is a need to develop techniques that exploit the structure of multimedia data in maximizing the efficiency of encryption algorithms.

There are three primary ways of delivering multimedia services to clients: *Unicast*, *broadcast* and *multicast*. Unicasting involves point-to-point communication between a server and a client device, broadcasting requires transmitting the same data to the entire client population, and multicasting is an efficient distribution mechanism from a source to a large group of clients. Protection of intellectual property (IP) has become a very critical issue in

* Corresponding author

digital multimedia distribution systems [3]. On the Internet (as well as terrestrial, satellite and cable systems), there are 3 essential requirements to prevent unauthorized access to copyrighted content:

- encryption of multimedia content
- cryptographic protection of content decryption keys (by encryption or other means)
- integrity of the critical data (copyright or usage rights) associated with the content.

In this paper, we do not address key management issues in digital multimedia content distribution. The interested reader is referred to two publications [1,4]. Our purpose is to survey the literature to understand the approaches for reducing the encryption load of a variety of client devices in distribution networks.

Selective encryption is currently an important research area. We will start with a classification and brief description of the proposed schemes in order to identify some of the related problems. This will be followed by a discussion of a number of future directions with potential improvements.

2. SELECTIVE ENCRYPTION

Presently, encryption appears to be the only technical tool that can be used to provide confidentiality in applications such as video conferencing, Pay TV and on-line video games. The cost of multimedia data compression can be aggravated by the additional need for protecting copyrighted digital content. It is argued that the encryption algorithms, which have been originally developed for text data, are not suitable for securing many real-time multimedia applications because of large data sizes. Software implementations of ciphers are usually too slow to process image and video data in commercial systems. Hardware implementations, on the other hand, add more cost to service providers and consumer electronics device manufacturers. A major recent trend is to minimize the computational requirements for secure multimedia distribution by “*selective encryption*” where only parts of the data are encrypted. A classification of the proposed schemes from the open literature is given in Table 1. To the best of our knowledge, such an extensive classification appears in the relevant literature for the first time.

Table 1. Classification of selective encryption schemes

Type of data	Domain	Proposal	Encryption Algorithm	What is encrypted?
Image	Frequency domain	Cheng & Li, 2000 [5]	No algorithm is specified.	Pixel and set related significance information in the two highest pyramid levels of SPIHT
		Droogenbroeck & Benedett, 2002 [6]	DES, Triple DES and IDEA	Bits that indicate the sign and magnitude of the non-zero DCT coefficients
		Pommer & Uhl, 2003 [7]	AES	Subband decomposition structure
	Spatial domain	Cheng & Li, 2000 [5]	No algorithm is specified.	Quadtree structure
		Droogenbroeck & Benedett, 2002 [6]	xor	Least significant bitplanes
		Podesser, Schmidt & Uhl, 2002 [8]	AES	Most significant bitplanes
Video	Frequency domain	Meyer & Gadegast, 1995 [9]	DES, RSA	Headers, parts of I-blocks, all I-blocks, I-frames of the MPEG stream
		Spanos & Maples, 1995 [10]	DES	I-frames, sequence headers and ISO end code of the MPEG stream
		Tang, 1996 [11]	Permutation, DES	DCT coefficients
		Qiao & Nahrstedt, 1997 [12]	xor, permutation, IDEA	Every other bit of the MPEG bit stream
		Shi & Bhargava, 1998 [13]	xor	Sign bit of DCT coefficients
		Shi, Wang & Bhargava, 1999 [14]	IDEA	Sign bit of motion vectors
		Alattar, A-Regib and Al-Semari [15]	DES	Every n^{th} I-macroblock, headers of all the predicted macroblocks, header of every n^{th} predicted macroblock
		Cheng & Li, 2000 [5]	No algorithm is specified.	Pixel and set related significance information in the two highest pyramid levels of SPIHT in the residual error
		Zeng & Lei, 2002 [16]	Permutation, xor	Selective bit scrambling, block shuffling, block rotation of the transform coefficients (wavelet and JPEG) and JPEG motion vectors
	Spatial domain	Cheng & Li, 2000 [5]	No algorithm is specified.	Quadtree structure of motion vectors and quadtree structure of residual errors
	Entropy codec	Wu & Kuo, 2000 [17]; Wu & Kuo, 2001 [18]	Multiple Huffman tables, multiple state indices in the QM coder	Encryption of data by multiple Huffman coding tables and multiple state indices in the QM coder

The schemes listed in Table 1 are used to protect 5 types of bitstreams:

- Uncompressed images
- JPEG compressed stream (images)
- MPEG compressed stream (video)
- Wavelet-based compressed stream (images and video)
- Quadtree-based compressed stream (images and video)

Note that a majority of the schemes are developed for JPEG and MPEG compliant streams. We will give a brief description of the proposed schemes.

(a) Schemes for Still Images

Selective Encryption in Frequency Domain

- **Cheng & Li, 2000:** In general, wavelet compression algorithms based on zerotrees transmit the structure of the zerotree with the significant coefficients. The SPIHT algorithm, for example, transmits the significance of the coefficient sets that correspond to trees of coefficients. Among the many different types of bits generated by the SPIHT algorithm, the proposed partial encryption scheme encrypts only the significance information related to pixels or sets in the two highest pyramid levels in addition to the parameter n that determines the initial threshold.
- **Droogenbroeck & Benedett, 2002:** In JPEG compression, the Huffman coder aggregates zero coefficients into runs of zeros and uses symbols that combine the run of zeros with magnitude categories for the non-zero coefficients that terminate the runs. These symbols are assigned 8-bit code words by the Huffman coder. The code words precede the appended bits that specify the sign and magnitude of the non-zero coefficients. In the proposed scheme, the appended bits corresponding to a selected number of AC coefficients are encrypted. The DC coefficients are left unencrypted because, it is argued, they carry important visible information and are highly predictable.
- **Pommer & Uhl, 2003:** The encoder chooses different decomposition schemes with respect to the wavelet packet subband structure for each image that needs to be protected. Classical best basis selection algorithm is not appropriate to determine a useful wavelet packet basis as it results in trees that share common features for many images, leading to a potential security weakness. Instead, the generation of the decomposition tree is randomized using a pseudo random number generator (PRNG). The tree carrying the subband decomposition structure is then secured for transmission with AES encryption.

Selective Encryption in Spatial Domain

- **Cheng & Li, 2000:** Quadtree image compression produces two logical parts: the quadtree and the parameters describing each block in the tree. The only parameter used by the authors to describe each block is the average intensity. As each intensity corresponds to a leaf node in the quadtree, the block intensities are called the *leaf values*. In the proposed partial encryption scheme, only the quadtree structure is encrypted. It can be used for both lossy compression (where each leaf is represented by the same number of bits) and lossless compression (where the number of bits to represent each leaf is different). For the transmission of the leaf values, two orderings are introduced: Leaf Ordering I (*inorder* traversal of the quadtree) and Leaf Ordering II (the leaf values are encoded one level at a time from the highest level to the lowest level). For security reasons, Leaf Ordering I is not recommended for lossy or lossless compression while Leaf Ordering II is reportedly secure for both.
- **Droogenbroeck & Benedett, 2002:** The decomposition of a gray scale image into its 8 bitplanes shows that the highest bitplanes exhibit some similarities with the original image while the least significant bitplanes look random. To exploit this property, some of the least significant bitplanes are encrypted. It is observed that at least 4 or 5 bitplanes need to be encrypted before the degradation becomes visible.
- **Podesser, Schmidt & Uhl 2002:** The gray scale image is decomposed into its 8 bitplanes and the most significant bitplanes are encrypted. After a number of experiments, it is observed that (1) the encryption of the

most significant bitplane is not secure enough, (2) selectively encrypting 2 bitplanes is sufficient if severe alienation of the image data is acceptable, and (3) encryption of 4 bitplanes provides high confidentiality.

(b) Schemes for Audio/Visual Streams

Selective Encryption in Frequency Domain

- **Meyer & Gadegast, 1995:** A new bit-stream called SEC MPEG is a modified version of MPEG, and incorporates selective encryption and additional header information for bit-error recovery. Four levels of security are implemented: (1) encryption of the headers from the sequence layer down to the slice layer, (2) encryption of parts of the *I*-blocks, (3) encryption of *I* frames and all *I*-blocks, and (4) full encryption.
- **Maples & Spanos, 1995:** A new security mechanism called Aegis is presented. Aegis encrypts the *I* frames of all groups of frames in an MPEG video stream. The encryption of the *I* frames is justified by the fact that they have great significance in the decompression of an MPEG stream whereas *B* and *P* frames represent only translations of the picture information found in adjacent *I* frames. Aegis is also used to encrypt the MPEG video sequence header and the ISO end code (last 32 bits of the MPEG video stream) to conceal the identity of the bit stream.
- **Tang, 1996:** The basic idea is to use a random permutation list to replace the zig-zag order in mapping the 8x8 DCT block to a 1x64 vector. Six experiments were conducted with different variations of permutation: (1) The DC coefficient is mapped to the first element in the 1x64 vector; the 63 AC coefficients are randomly permuted. (2) The DC coefficient is set to zero; the 63 AC coefficients are permuted according to the zig-zag order. (3) All coefficients are randomly permuted and the DC coefficient is not in the first position of the vector. (4) All coefficients are permuted according to the zig-zag order; the last AC coefficient is set to zero. (5) The DC coefficient is split according to the splitting procedure; all coefficients are permuted randomly. The splitting procedure is defined as follows: Let $d_7d_6\dots d_1d_0$ be the 8-digit binary representation of a DC coefficient. It is split into two numbers $d_7d_6d_5d_4$ and $d_3d_2d_1d_0$, which are both in the range of [0,15]. Set the value of the DC coefficient to be $d_3d_2d_1d_0$, and the value of the last AC coefficient to be $d_7d_6d_5d_4$. (6) The DC coefficient is encrypted by the function $f_i(x_{i1}, \dots, x_{i8}) = (k \oplus (x_{i1} \dots x_{i8} \dots x_{i81} \dots x_{i88}))_{8^{*i+1}, \dots, 8^{*i+8}}$, where k is a 64-bit secret key, x_{ij} is the ij -th bit of the binary sequence formed by grouping eight DC coefficients together, and \oplus is the binary *xor* operation; splitting and random permutation procedures are applied.
- **Qiao & Nahrstedt, 1997:** The basic approach is to take the chunk $a_1a_2a_3\dots a_{2n-1}a_{2n}$ of an *I*-frame and create the two byte streams $a_1a_3\dots a_{2n-1}$ (odd list) and $a_2a_4\dots a_{2n}$ (even list). The substreams are then *xored* to obtain the ciphertext $c_1c_2c_3\dots c_n$ which is concatenated to $E(a_2a_4\dots a_{2n})$, where E denotes an encryption function. If $a_2a_4\dots a_{2n}$ has no repeated pattern, the secrecy depends on the function E as $a_2a_4\dots a_{2n}$ can be considered to be a one-time pad. Using the basic idea, an algorithm is developed with several keys: *KeyM* is used to derive the two byte streams; each *Key_i* ($i=1, \dots, 8$) shuffles a chunk of data to obtain a non-repeated pattern with a length of 1/2 frame (it was observed that the non-repeated patterns have a life time of over only one 1/16 chunk); *KeyF* is assigned to each frame to change the pattern of choosing even and odd lists (which is applied to *KeyM* repeatedly and to *Key_i*'s to derive new keys for each frame); and *KeyE* is the encryption key for the function E . *KeyM*, *Key_i*'s and *KeyF*'s can be encrypted with *KeyE*, and *KeyE* can be sent via a separate secure channel. Alternatively, *KeyM* and *Key_i*'s can be sent in a separate secure channel.
- **Shi & Bhargava, 1998:** The Video Encryption Algorithm (VEA) uses a secret key to randomly change the sign bits of the DCT coefficients of MPEG video. VEA's secret key $k = b_1b_2\dots b_m$ is a randomly generated bitstream of length m . If the sign bits of DC and AC coefficients are represented by $S = \dots s_1 \dots s_m s_{m+1} \dots s_{2m}$, VEA's encryption function is $E_k(S) = \dots (b_1 \oplus s_1) \dots (b_m \oplus s_m)(b_1 \oplus s_{m+1}) \dots (b_m \oplus s_{2m}) \dots$, where \oplus is the binary *xor* operation. VEA does not have a limit on the key length or number of keys. Multiple keys can be used in several ways: 2 keys, one for Y blocks, one for Cb and Cr blocks; 3 keys, one for Y blocks, one for Cb blocks and one for Cr blocks; 3 keys, one for *I* frames, one for *B* frames and one for *P* frames.
- **Shi, Wang & Bhargava, 1999:** Real-time Video Encryption Algorithm (RVEA) is based on the previous work VEA and introduces two improvements: it adopts ciphers to increase security and limits the maximum number of selected bits to bound the computation time. For each 16x16 macroblock in a video slice, RVEA selects at most 64 sign bits. The order of the sign bits in the six 8x8 blocks Y1, Y2, Y3, Y4, Cr and Cb is defined in a specific way with the consideration that DC coefficients are more significant than AC coefficients and lower

frequency AC coefficients are more significant than higher frequency AC coefficients. The selected sign bits are encrypted with DES or IDEA and put back in their original positions.

- **Alattar, A-Regib and Al-Semari, 1999:** Three methods are proposed to improve the performance of an earlier work by two of the co-authors. In the first method, encryption is applied to the data associated with every n^{th} I -macroblock. In the second method, the headers of all predicted macroblocks are encrypted together with the data associated with every n^{th} I -macroblock. To reduce the computational load, the third method encrypts only the headers of every n^{th} predicted macroblock in addition to the data associated with every n^{th} I -macroblock.
- **Cheng & Li, 2000:** An extension of the image compression algorithm called the Set Partitioning in Hierarchical Trees (SPIHT) algorithm (which is an implementation of zerotree wavelet image compression) is used to encode the residual error. Since the residual error is an image frame, the partial encryption scheme for wavelet image compression can be directly applied to residual error coding. It is observed that the relative size of the important part of the residual error is similar to the size in image compression.
- **Zeng & Lei, 2002:** The frequency domain scrambling technique divides the transform coefficients into blocks/segments and performs some or all of the following three operations: selective bit scrambling, block shuffling and block rotation of the transform coefficients and motion vectors. Two compression schemes are used to illustrate the approach: wavelet transform based compression and 8x8 DCT based compression. In the simulations, several combinations of the three scrambling operations are tested in the wavelet and DCT domains.

Selective Encryption in Spatial Domain

- **Cheng & Li, 2000:** Video compression algorithms generally address motion compensation and residual error coding. An extension of the quadtree compression algorithm is used to encode both motion vectors and residual errors. Since the residual error is an image frame, partial encryption scheme for quadtree image compression can be directly applied to residual error coding. It is observed that the relative size of the important part of the residual error is similar to the size in image compression. In encoding motion vectors, the leaf values become the motion vectors. As in the original scheme for images, the encrypted part is the quadtree decomposition, not the motion vectors. For low-resolution videos where the maximum height of the tree may be small enough to allow exhaustive tree enumeration, the motion vectors are completely encrypted.

Entropy Codec Design

- **Wu & Kuo, 2000; Wu & Kuo, 2001:** Base on a number of observations, the authors argue that selective encryption (i.e., the selection of the most important coefficients from a compression system and their encryption with a conventional cipher) may not be effective in two situations: (1) The media compression system is based on an orthogonal transform followed by quantization, and (2) The media compression system contains entropy coding at the last stage. Using multiple statistical models, they investigate a different encryption methodology that turns entropy coders into ciphers. This methodology allows the construction of two selective encryption schemes by application to the Huffman coder and the QM coder, two of the most popular entropy coders in multimedia compression. Both coders have very simple statistical models; the model of the Huffman coder is usually a fixed-size non-adaptive binary tree, and the initial state of the QM coder includes only three integer numbers. Since hiding the Huffman coding table or the initial state of the QM coder does not provide sufficient secrecy, it is argued that the problem of a limited key/model space can be overcome by using m statistical models instead of only one. The first of the proposed encryption schemes makes use of multiple Huffman tables and the second multiple indices in the QM coder estimation state machine.

CURRENT ISSUES

The above proposals suffer one or more of the following problems that may have a serious negative impact on a given application:

1. Insufficient security
2. Decrease in the compression performance of entropy coding
3. Insignificant computational reduction with respect to total encryption
4. Lack of bitstream compliance
5. Increase in key size

- *Insufficient security:* Restricting the encryption to the I frames [10] does not make the information in the P and B frames useless if their base frames are correlated [19]. The Zig-Zag Permutation Algorithm [11] is vulnerable to known-plaintext and ciphertext-only attacks [20]. Encryption of the sign bit (and also more significant bits) of every DCT coefficient does not work because useful image content can be recovered by assigning all DC coefficients to 128 and all AC coefficients to positive [17]. To create visible degradation, a minimum number of 4 or 5 least significant bitplanes should be encrypted in uncompressed images [6]. Two types of ciphertext only attacks on bitplane encryption (replacement attack and reconstruction attack) show that encryption of the most significant bitplane is not secure enough; at least 4 bitplanes need to be protected [8]. In [7], the amount of data to be encrypted for a given image is extremely small as no image data (e.g., transform coefficients) needs to be protected. Simulation results in [15] show that it is sufficient to encrypt every other I -macroblock and the header of every other predicted macroblock to disguise the video completely. When this simulation is repeated for $n=3$, there is a clear degradation in the security level. The motion becomes easily recognizable when the encrypted MPEG video sequences are played back.
- *Decrease in the compression performance of entropy coding:* In MPEG compression, the non-zero AC coefficients are generally concentrated in the upper left corner of the 8×8 block. The application of the zig-zag ordering to map the 64 coefficients into a vector and entropy encoding result in an efficient compression. Any other ordering to map the coefficients [11] will decrease the compression ratio. The tests conducted to compare the sizes resulting from zig-zag ordering and non-zig-zag ordering show that the size increase can be as much as 46% [20]. In the wavelet based systems, block shuffling or block rotation alone introduces up to 5% bit rate increase for the same PSNR, and in the 8×8 DCT based systems, shuffling along slices with/without sign encryption increases the bit rate by about 20% for the same PSNR [16]. The multiple Huffman coding tables generated in entropy codec design [17,18] should not lead to a relative decrease in the compression ratio. The overhead to achieve this goal is to generate each Huffman table from a different set of training images (or audio pieces). The random generation of decomposition trees may not result in optimal compression efficiency when compared with wavelet packet coders targeted for image compression [7].
- *Insignificant computational reduction with respect to total encryption:* Encrypting only the I frames [10] can reduce the cryptographic computations by 50%. An increase in the frequency of I frames will increase the security level but it will also increase both the length of the stream and encryption/decryption time [20]. Encryption of all I -blocks [9] raises several problems: Identification of I -blocks in a P or B frame introduces the overhead of searching the MPEG stream; some MPEG streams contain I frames only, reducing the selective algorithm to full encryption; the number of I -blocks in P or B frames can be of the same order as the number of I -blocks in I frames [12]. The video encryption algorithm [12] is based on encrypting every other bit of the MPEG stream. Since one half of the bits has to be encrypted, the required computation is 50% of total encryption without considering the overhead of the selection process [17]. Encryption of at least 4 or 5 of the least significant bitplanes [6] means that the algorithm's computational cost is at least 50% to 60%. Encryption of 4 most significant bitplanes [8] will result in a computational cost of 50%.
- *Lack of bitstream compliance:* As SEC MPEG [9] requires additions and changes to the standard MPEG bitstream headers, a SEC MPEG bitstream is not compatible with the MPEG standard. To view unencrypted SEC MPEG bitstreams, a special encoder and decoder would be needed. The quadtree image compression [5] is not part of any common image compression standards. Hence, the proposed partial encryption scheme may have limited use in commercial applications.
- *Increase in key size:* Several keys are used in the Video Encryption Algorithm [12]. A 128-bit or 256-bit $KeyM$; 8 Key_i 's, each with a length of 160 bits; a 64-bit $KeyF$ for each frame, and finally $KeyE$ for the encryption function. In the entropy codec design [17,18], multiple statistical models are used. In Huffman codec design, 2^k different Huffman tables numbered from 0 to $(2^k - 1)$ and a random vector $P = (p_1, p_2, \dots, p_n)$, where each p_i is a k -bit integer varying from 0 to $(2^k - 1)$, are generated. The i^{th} symbol in the original data stream is encoded with table $p_{(i-1 \pmod n)+1}$. The 4 indices employed in the QM coder are set to hide initial values and used alternately in a secret order to encode the input bitstream. First, a random key $K = \{(s_0, s_1, s_2, s_3), (p_0, \dots, p_{n-1}), (o_0, \dots, o_{n-1})\}$ is generated, where s_i is a 4-bit integer and p_i and o_i are 2-bit integers. After an initialization of the 4 state indices (I_0, I_1, I_2, I_3) to (s_0, s_1, s_2, s_3) , to encode the i^{th} bit in the original stream, the index $I_{p(i \pmod n)}$ (called the active index) is used to determine the probability estimation Q_e . If state update is required after encoding the i^{th} bit in the input stream, all state indices except $I_{o(i \pmod n)}$ are updated. In the frequency domain selective scrambling of digital video [16], depending on the transform type, several keys are needed for

selective bit scrambling, block shuffling and block rotation of the transform coefficients and motion vectors. In wavelet based systems, for example, the sign of each coefficient can be changed using a key-based cryptographically secure pseudo random process. In block shuffling, coefficients are shuffled according to a shuffling table generated by a key. In block rotation, each block of coefficients is rotated to form an encrypted block, where the encrypted block is selected from a set of 8 blocks that are rotated versions of the original block.

Two publications evaluate the performance of some of the above proposals, and discuss the tradeoffs among several metrics such as security level, encryption speed and compression efficiency [19,20].

As noted earlier, most of the modern encryption algorithms were developed for text data, i.e., the simplest form of multimedia. The more complex forms (especially video) may vary substantially in their cost of creation as well as their storage and communication requirements. When compared with bank account information, for example, the value of a movie is much lower while the bit rate is much higher.

There is a wide spectrum of secure multimedia applications with different requirements. They range from military applications that mandate total data obscurity to applications where a part of the multimedia data needs to be visible to allow searching in a shared database. Hence, it is desired to develop an encryption-based technology that is appropriate for many of such scenarios. The desirable attributes of such a technology include [6,12,16,17,18,20,21]:

- it should provide sufficient security for a range of multimedia applications,
- it should preserve the size of the original unencrypted bitstream,
- it should result in substantial computational reduction with respect to total encryption,
- it should produce a bitstream that is compliant to the standard formats,
- it should not create a key whose size is much longer than those of commonly used modern ciphers, and
- it should identify the portions of the multimedia data to be encrypted.

FUTURE DIRECTIONS

We believe that the following areas of research are promising in determining the future of selective encryption of multimedia content:

- As key management (i.e., the generation, storage and replacement of keys) is a critical issue in all encryption based security systems, it cannot be separated from the design of secure multimedia distribution. In most distribution architectures, multimedia content is encrypted with a symmetric key which also needs to be protected in transmission to the receiver. A common tool of achieving the protection of the decryption key is public-key cryptography. The difficulty in cryptanalizing public-key ciphers would provide reasonable security in most applications. In some of the selective encryption schemes surveyed in this paper, key generation is not properly addressed. To be able to maintain entropy coding efficiency, these proposals avoid using standard ciphers, and instead employ key-based permutation, resulting in key lengths much longer than 64 or 128 bits needed for cryptographically strong symmetric ciphers. Furthermore, the security of operations such as coefficient shuffling may not be as strong as that of hybrid modern ciphers that employ both permutation and substitution. Hence, the storage and security requirements of key management need to be discussed in greater detail in future proposals.
- The classification in Table 1 shows that a common approach in selective video encryption is to integrate compression and encryption processes whereby only a subset of the transform coefficients (or some of their bits) are scrambled. However, it has been observed that energy concentration via transform domain compression does not often imply intelligibility concentration [18]. An important consequence of this observation is that selective encryption is not effective for orthogonal transform based compression followed by entropy coding (today's compression standards - JPEG and JPEG 2000 for image compression, MPEG for video compression, and MP3 for audio compression). It may therefore be desirable to keep compression and encryption as two separate processes. How, then, does intelligibility concentration map to energy concentration? First, one major focus should be on the new compression standards such as JPEG 2000, the wavelet transform based image coding system, and MPEG-4, the object-based video coding system with target applications such as Internet Multimedia, Interactive Video Games, Interpersonal Communications

(Videoconferencing, Videophone etc.), Networked Database Services, Interactive Storage Media (optical disks, etc.), and Wireless Multimedia. Scalable video compression[†] can be another promising area of research as it enables the video codec to adapt to different client capabilities and network conditions in heterogeneous environments. Development of a general selective encryption scheme for still images and video with the listed desirable attributes is a complex problem. Ideally, our findings for still images should be extendable to video protection with more computational requirements. Identification of the most significant parts of multimedia data is a key step for selective encryption.

- Content and service providers are developing new business models that specify how digital copyrighted content can be consumed. The usage rights need to be delivered to the consumers together with the content and the decryption keys. The simplest form of this data is the *Copy Control Information (CCI)* that expresses the conditions under which a consumer is allowed to make a copy of a content legally accessed. An important subset of CCI is the two Copy Generation Management System (CGMS) bits for digital copy control: “11” (copy-never), “10” (copy-once), “01” (no-more-copies), and “00” (copy-free). It is possible to associate the CCI with the content in two ways: (1) the CCI is included in a designated field in the A/V transport stream, and (2) the CCI is embedded as a watermark into the A/V stream. As the business models evolve, there will be a need to carry more information with the protected content. Development of robust watermarking systems with sufficient capacity is a very challenging open research issue [21]. A watermarking proposal for copyright protection needs to be resistant to both intentional attacks and normal A/V data processing (compression, A-D-A conversion, scaling, filtering, rotation, cropping, etc.). An interesting idea would be to integrate watermarking and encryption processes. For example, a particular technique to deliver a piece of the decryption key is secret sharing. It is possible to use a *prepositioned secret sharing scheme* [22] where an “*activating share*” is delivered to the receivers and the secret key is reconstructed using the activating share and prepositioned information (i.e., the shares stored in the receiver) [23]. The scheme has been proposed to transport decryption keys in unicast, broadcast and multicast architectures [24,25]. The Internet is a boundless network with a wide range of channel bandwidths and client device capabilities. It may therefore be necessary to add a feature to the proposed selective schemes where the selection criteria can be chosen dynamically as the content is being distributed. Putting the selection criteria in the activating share allows the construction of a *dynamic* selective encryption scheme where the selection criteria can be changed as needed per application or even per image type in an application. Furthermore, prepositioned secret sharing can also be used to develop dynamic watermarking schemes.

CONCLUSIONS

Selective encryption is the process of encrypting only parts of a multimedia content to reduce the computational requirements of client devices in real-time applications. We presented a survey and classification of the proposed schemes in the open literature. The major problems associated with most of these schemes are insufficient security, decrease in the compression performance of entropy encoding, insignificant computational reduction with respect to total encryption, lack of bitstream compliance and increase in encryption key size. Promising future directions of research include more emphasis on key management, resolving the conflict between compression and encryption, an investigation of robust watermarking techniques to carry usage rights information, and finding ways to change the selection criteria dynamically. With the increasing availability of digital distribution and storage technologies, the demand for multimedia services is on the rise. We hope that efficient solutions for multimedia security will lead to flourishing businesses to make the content owners, service providers, CE device manufacturers and the consumers happy.

[†] Scalable video compression is the encoding of a single video stream in different layers, each layer with its own bit rate.

REFERENCES

- [1] Eskicioglu, A. M., Town, J. and Delp, E. J., "Security of Digital Entertainment Content from Creation to Consumption," *Signal Processing: Image Communication, Special Issue on Image Security*, 18(4), April 2003, pp. 237-262.
- [2] Windows Media Rights Manager, available at <http://www.microsoft.com/windowsmedia>
- [3] Eskicioglu, A. M. and Delp, E. J., "Overview of Multimedia Content Protection in Consumer Electronics Devices," *Signal Processing: Image Communication*, 16(5), April 2001, pp. 681-699.
- [4] Eskicioglu, A. M., "Multimedia Security in Group Communications: Recent Progress in Key Management, Authentication, and Watermarking," to appear in *ACM Multimedia Systems Journal, Special Issue on Multimedia Security* in 2003.
- [5] Cheng, H. and Li, X., "Partial Encryption of Compressed Images and Video," *IEEE Transactions on Signal Processing*, 48(8), 2000, pp. 2439-2451.
- [6] Van Droogenbroeck, M. and Benedett, R., "Techniques for a Selective Encryption of Uncompressed and Compressed Images," *Proceedings of Advanced Concepts for Intelligent Vision Systems (ACIVS) 2002*, Ghent, Belgium, September 9-11, 2002.
- [7] Pommer, A. and Uhl, A., "Selective Encryption of Wavelet-Packet Encoded Image Data," to appear in *ACM Multimedia Systems Journal, Special Issue on Multimedia Security* in 2003.
- [8] Podesser, M., Schmidt, H.-P. and Uhl, A., "Selective Bitplane Encryption for Secure Transmission of Image Data in Mobile Environments," *5th Nordic Signal Processing Symposium*, on board Hurtigruten, Norway, October 4-7, 2002.
- [9] Meyer, J. and Gadegast, F., "Security Mechanisms for Multimedia Data with the Example MPEG-1 Video," *Project Description of SEC MPEG*, Technical University of Berlin, Germany, May 1995.
- [10] Spanos, G. A. and Maples, T. B., "Performance Study of a Selective Encryption Scheme for the Security of Networked, Real-time Video," *Proceedings of 4th International Conference on Computer Communications and Networks*, Las Vegas, NV, September 20-23, 1995.
- [11] Tang, L., "Methods for Encrypting and Decrypting MPEG Video Data Efficiently," *Proceedings of the 4th ACM International Multimedia Conference*, Boston, MA, November 18-22, 1996, pp. 219-230.
- [12] Qiao, L. and Nahrstedt, K., "A New Algorithm for MPEG Video Encryption," *Proceedings of the 1st International Conference on Imaging Science, Systems and Technology (CISST '97)*, Las Vegas, NV, July 1997, pp. 21-29.
- [13] Shi, C. and Bhargava, B., "A Fast MPEG Video Encryption Algorithm," *Proceedings of the 6th International Multimedia Conference*, Bristol, UK, September 12-16, 1998.
- [14] Shi, C., Wang, S.-Y. and Bhargava, B., "MPEG Video Encryption in Real-Time Using Secret key Cryptography," *1999 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'99)*, Las Vegas, NV, June 28 - July 1, 1999.
- [15] Alattar, A. M., Al-Regib, G. I. and Al-Semari, S. A., "Improved Selective Encryption techniques for Secure Transmission of MPEG Video Bit-Streams," *Proceedings of the 1999 International Conference on Image Processing (ICIP '99)*, Vol. 4, pp. 256-260, Kobe, Japan, October 24-28, 1999.
- [16] Zeng, W. and Lei, S., "Efficient Frequency Domain Selective Scrambling of Digital Video," *IEEE Transactions on Multimedia*, 2002.
- [17] Wu, C.-P. and Kuo, C.-C. J., "Fast Encryption Methods for Audiovisual Data Confidentiality," *SPIE International Symposia on Information Technologies 2000*, Boston, MA, November 2000, pp. 284-295.
- [18] Wu, C.-P. and Kuo, C.-C. J., "Efficient Multimedia Encryption via Entropy Codec Design," *Proceedings of SPIE Security and Watermarking of Multimedia Content III*, Volume 4314, San Jose, CA, January 2001.

-
- [19] Agi, I. and Gong, L., "An Empirical Study of Secure MPEG Video Transmission," *Proceedings of the Internet Society Symposium on Network and Distributed System Security*, San Diego, CA, February 1996, pp. 137-144.
- [20] Qiao, L. and Nahrstedt, K., "Comparison of MPEG Encryption Algorithms," *International Journal on Computer and Graphics, Special Issue on Data Security in Image Communication and Network*, 22(3), 1998.
- [21] Doerr, G. and Dugelay, J.-L., "A Guide Tour of Video Watermarking," *Signal Processing: Image Communication*, 18(4), 2003, pp. 262-382.
- [22] Simmons, G. J., "Prepositioned shared secret and/or shared control schemes," *Advances in Cryptology – EUROCRYPT '89 Proceedings*, Springer-Verlag, 1990, pp. 436-467.
- [23] Eskicioglu, A. M., Delp, E. J. and Eskicioglu, M. R., "New Channels for Carrying Copyright and Usage Rights Data in Digital Multimedia Distribution," to be presented at the *International Conference on Information Technology: Research and Education*, Newark, NJ, August 10-13, 2003.
- [24] Eskicioglu, A. M. and Delp, E. J., "A Key Transport Protocol Based on Secret Sharing – Applications to Information Security," *IEEE Transactions on Consumer Electronics*, Vol. 48. No. 4, November 2002, pp. 816-824.
- [25] Eskicioglu, A. M. and Eskicioglu, M. R., "Multicast Security Using Key Graphs and Secret Sharing," *Proceedings of the Joint International Conference on Wireless LANs and Home Networks (ICWLHN 2002) and Networking (ICN 2002)*, Atlanta, GA, August 26-29, 2002, pp. 228-241.