

Chapter 2

INTRODUCTION TO MULTIMEDIA ENCRYPTION

2.1 MULTIMEDIA AND CRYPTOGRAPHY

In the recent years, there has been a tremendous improvement and emergence of technologies for communications, coding, and retrieval of digital multimedia. Such an environment has allowed for the realization of many fascinating multimedia applications related to nearly all aspects of life. Almost instantaneous delivery of entertainment videos, pictures and music is available to everyone who is connected to a multimedia distribution system. Businesses and other organizations are now able to perform real-time audioconferencing and videoconferencing, even over a non-dedicated channel. By connecting to a medical imaging database system, the experts, even the ones from remote areas, can instantaneously receive and review relevant medical images using the power of image coding and image retrieval techniques. However, many multimedia distribution networks are open public channels and, as such, are highly insecure. An eavesdropper can conveniently intercept and capture the sensitive and valuable multimedia content traveling in a public channel. Fortunately, the magic art of cryptography can help prevent this.

In general, multimedia security is achieved by a method or a set of methods used to protect the multimedia content against unauthorized access or against unauthorized distribution. These methods are heavily based on cryptography and they provide either communication security, or security against piracy (DRM and watermarking). Communication security of multimedia data can be accomplished by means of conventional cryptography. Multimedia data could be entirely encrypted using a cryptosystem such as DES [9], IDEA [10] or AES [11]. In many cases, when the multimedia is textual or static data, and not a real-time streaming media, we can treat it as an ordinary binary data and use the conventional

encryption techniques. Encrypting the entire multimedia stream using standard encryption methods is often referred to as the *naïve approach*. The naïve approach is usually suitable for text, and sometimes for small bitrate audio, image, and video files that are being sent over a fast dedicated channel. Secure Real-time Transport Protocol [1], or shortly SRTP, is an application of the naïve approach. In SRTP, multimedia data is packetized and each packet is individually encrypted using AES. The naïve approach enables the same level of security as that of the utilized conventional cryptosystem. Unfortunately, encrypting the entire bitstream is typically not possible for higher bitrate multimedia, especially when the transmission is done over a non-dedicated channel.

Due to a variety of constraints, communication security for streaming multimedia is harder to accomplish [6,7,8]. Such constraints include real-time processing, non-dedicated channels with limited or varying bandwidth, high multimedia bitrate, and more. Thus, a communication encryption of many video and audio multimedia is not simply the application of established conventional encryption algorithms to their binary sequence. It involves careful analysis to determine and identify the optimal encryption method. Current research is focused towards exploiting the format specific properties of many standard multimedia formats in order to achieve the desired performance. This is referred to as the *selective encryption*. This type of encryption is obviously preferred when compression and decompression algorithms can hardly keep up with the required bitrate, even when these algorithms are accelerated by a dedicated hardware. In some cases, encryption and decryption algorithms could also be accelerated by hardware. However, software implementations are often preferred due to their flexibility and low cost. Figure 2.1 shows the logical stages that are taken during both the naïve approach and the selective approach.

The naïve approach is obviously more straightforward to implement, but the entire bitstream is passing through the computationally expensive encryption and decryption stages. Although more complex to implement, selective encryption will only process the selected bits through encryption and decryption stages. More recently, full encryption approaches based on chaotic maps have been proposed for securing the multimedia.

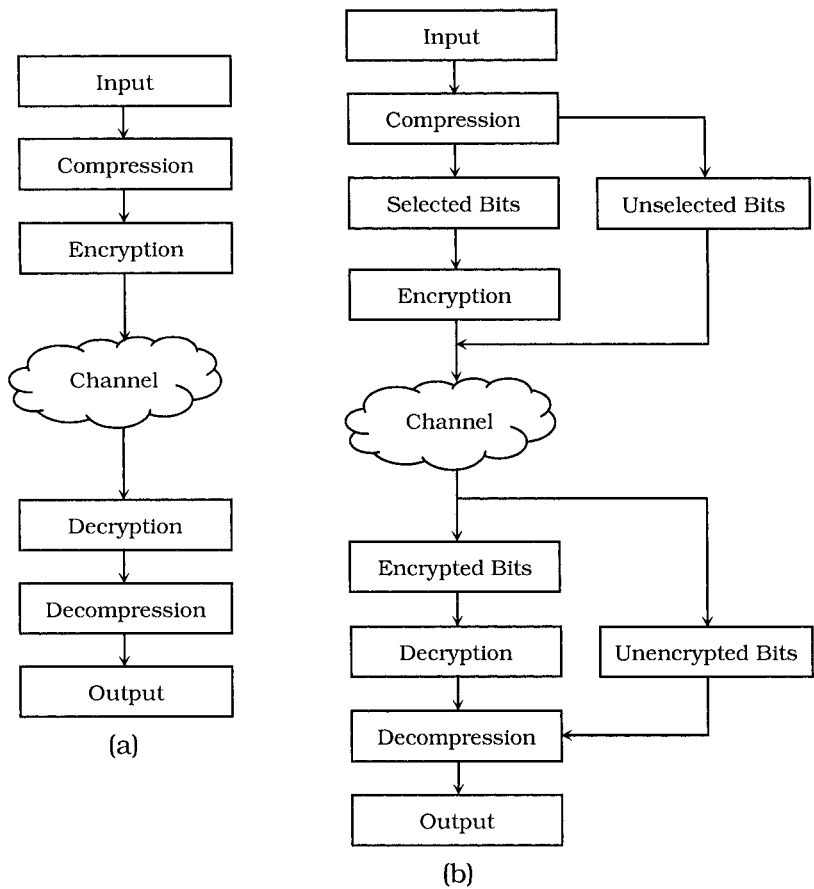


Figure 2.1. Stages taken during multimedia encryption using (a) the naïve approach, and (b) the selective approach.

Chaos-based cryptography had been studied for quite some time, but the encryption approaches using 1-D, 2-D and even 3-D chaotic maps specifically designed for encrypting digital images and videos had been only recently proposed. Chaos-based multimedia encryption algorithms are of high interest due to their extremely fast performances and suitability for multimedia data.

We have already mentioned several reasons as to why we need multimedia-specific cryptosystems. Most of the traditional cryptosystems were designed specifically to secure and encrypt text messages. Digital multimedia data, however, have much different

properties, which make them less suitable for many conventional cryptosystems. Multimedia data includes images, videos, speech, and audio, which are usually very large-sized and bulky. When encrypting such bulky data, the performance overhead that is introduced with some of the conventional ciphers is generally too expensive for real-time applications. Secondly, an extremely high data redundancy is often present in many uncompressed images and videos. Consequently, some of the conventional block ciphers may fail to obscure all visible information unless advanced modes of operation are used. Figure 2.2 shows how AES that operates in its basic ECB mode fails to disguise the image. As one can see, if the more complex modes are used, AES is more effective. However, modes that are more computationally complex consequently run slower. Different modes of block ciphers will be discussed further in the next chapter.

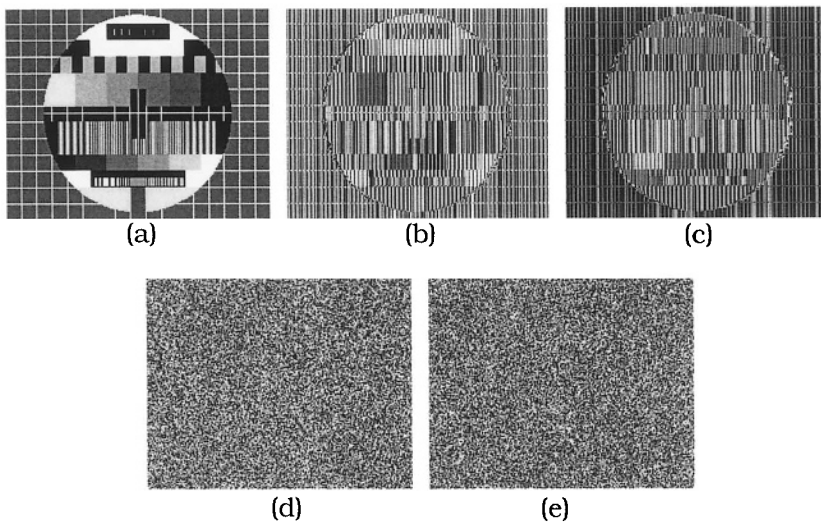


Figure 2.2. Encrypting a redundancy rich uncompressed image: (a) the original image, (b) the image encrypted with AES-128 in ECB mode, (c) the image encrypted with AES-256 in ECB mode, (d) the image encrypted with AES-128 in CBC mode, and (e) the image encrypted with AES-256 in CBC mode.

In the case of digital videos, consecutive frames are extremely similar and most likely only few pixels (if any) would differ from frame to frame. The conventional ciphers are designed with a good avalanche

property, however, in the case of videos, an extremely strong avalanche property is required. Furthermore, chosen/known-plaintext attack can be easily performed for videos, and a cipher that is extremely robust to chosen/known-plaintext attacks is desired. Another drawback of conventional cryptosystems relates to the compressed multimedia. Multimedia compression and encryption are usually very incompatible. Encrypting the multimedia content before compression removes a lot of redundancy and this results in a very poor compression ratio. On the other hand, encrypting the data after compression destroys the codec format, which causes decoders to crash. Finally, for many applications, we would like to have very light encryption that preserves some perceptual information. This is impossible to achieve with conventional cryptosystems alone, which most likely degrade the data to a perceptually unrecognizable content. Having cryptosystems that preserve a rough view of the high-quality media material is preferred for many real-life applications, such as video pay-per-view system in which a degraded but visible content could potentially influence a consumer to order certain paid services.

2.2 SECURITY VS. PERFORMANCE

Undoubtly, implementing security based on a cryptographic protocol introduces a performance overhead during the multimedia processing. The size of an overhead depends on many factors. In most cases, utilizing only the general software or compiler based optimization techniques yields only modest improvements. The complexity of an encryption/decryption algorithm is one of the major factors affecting the secure multimedia system performance. Clearly, a fast, yet secure algorithm is desired. In fact, the main goal of multimedia encryption is to significantly reduce the performance overhead while maintaining the desired level of security. Only then can one expect the secure real-time delivery of high-quality and large bitrate multimedia over a non-dedicated channel. Unfortunately, this is not easy to accomplish.

Researchers and experts put a lot of effort in designing good multimedia encryption algorithms. As a result, there is a significant number of proposals that rely on the selective encryption. Once a selection of bits to be encrypted is made, a conventional encryption could be used to secure them. By a “conventional cryptosystem” we mean a cryptosystem that is either one of the encryption standards

(DES or AES), or a well-established cryptosystem that was designed by the cryptography experts and that has been around for a while. Many cryptographic systems fulfill the aforementioned credibility requirements. Unfortunately, there are far more of them that do not meet these requirements. It is often dangerous to blindly trust a cryptosystem that had been recently proposed, as the cryptographic community needs some time to thoroughly investigate its security and possibilities of an attack. In this book, we will describe few conventional cryptosystems that can be safely used as a basis for many multimedia selective encryption approaches. These cryptosystems withstood many years of attacks and we can safely assume that their security is likely to remain strong. Modifying or simplifying conventional cryptosystems in order to improve the multimedia encryption performance is usually a bad idea. It is as bad an idea as if a construction worker would suddenly decide to build a wall using the cardboard boxes instead of bricks to save the cost or to speed up the construction process. Before you know it, someone would almost effortlessly break the wall, just as someone would almost effortlessly break the “simplified” cryptosystem. Yet, a number of early multimedia security solutions used far oversimplified encryption algorithms in order to produce faster performances. This kind of cryptography does not provide security. It hardly provides a temporary inconvenience for the attacker.

This book, however, is not about cryptography. Cryptography is complex, difficult, and a highly mathematically involved discipline. This book is about applying cryptography to achieve the desired multimedia security and protection. While some major concepts of cryptography will be discussed to provide a better understanding of topics ahead, the primary focus of this part of the book is on the multimedia encryption techniques, their performance, and their security.

Although slightly outdated, “Handbook of Applied Cryptography” by van Oorschot, Menezes, and Vanstone [2] is still a top technical reference to the subject of general cryptography, while Bruce Schneier’s “Applied Cryptography” [3] is still an excellent resource that in addition includes the C code for many important cryptography-related algorithms. An excellent new book on the subject is also “Fundamentals of Computer Security” by Pieprzyk, Hardjono, and Seberry [4]. Finally, Doug Stinson’s “Cryptography: Theory and Practice” [5] is one of the best textbooks about

cryptography that covers the core material and includes numerous examples and exercises.

2.3 LEVELS OF SECURITY

A given cryptosystem provides a certain level of security. Most available cryptographic systems are fully or partially scalable, in the sense that one can choose different security levels. Scalability is usually achieved by allowing variable key sizes or by allowing different number of iterations, or rounds. A higher level of security is achieved with larger key sizes or larger number of rounds. Consequently, an algorithm becomes more complex and slower. Therefore, it is important to carefully assess the value and importance of the multimedia content that needs to be encrypted. According to the content value, the cryptosystem can be properly scaled to allow an optimal performance-security ratio.

There is yet another important consideration in the case of multimedia encryption. Applying selective encryption has its own scalability chart. Obviously, selecting all bits for encryption is inefficient, but certainly the most secure choice. The amount of selected bits and type of selected bits uniquely determine the complexity of an attack. This type of scalability is much harder to grade, and the tools for its assessment are not yet fully developed nor established. The reason for this is that the conventional, non-multimedia encryption algorithms have a long history. Analytical methods and tools were developed over a course of many years, and in many instances rigorous mathematical proofs allowed for proper evaluation. On the other hand, multimedia specific encryption algorithms are relatively new. First such algorithms started to appear in the mid-1990s, and in most cases, solid methods for rigorous security analysis of such algorithms are simply not available. An additional cause for this is the high complexity of many encoded multimedia formats. For example, some of the available video codecs produce such intricate data that their true mutual correlation is extremely complex to understand. To evaluate how much the non-encrypted data can reveal about the encrypted bits is consequently a hard problem.

The problem of rigorous security analysis of the proposed multimedia encryption algorithms is interdisciplinary in nature. It requires proficiency in multimedia processing, as well as a

proficiency in cryptography and cryptanalysis. In this book we try to overcome these problems as much as possible, by providing careful analysis of the proposed image, video, audio, and speech encryption algorithms. To our knowledge, that is what makes this book unique.

2.4 DEGRADATION AND SCRAMBLING

Deciding upon what level of security is needed is harder than it looks. To identify an optimal security level, we have to carefully compare the cost of the multimedia information to be protected versus the cost of the protection itself. If the multimedia to be protected is not that valuable in the first place, it is sufficient to choose a relatively light level of encryption. On the other hand, if the multimedia content is highly valuable or represents industrial, governmental or military secrets, the cryptographic security level must be the highest possible.

For many real-world applications (such as pay-per-view) the content data rate is typically very high, but the monetary value of the content may not be high at all. Thus, very expensive attacks are not attractive to adversaries, and lightweight encryption, often called *degradation*, may be sufficient for distributing such multimedia content. For such applications, DRM is of much more interest since the content owner and the content provider both seek to enforce the copyrights and the distribution rights. When degradation is applied to a multimedia content, the content is usually still perceptible to some degree. For instance, one may still see the objects in a degraded video, but the visual quality should be unacceptable for entertainment purposes (see Figure 2.3-b). In contrast, applications such as videoconferencing or videophone (or even Internet phone) may require a much higher level of confidentiality. If the videoconference is discussing important industrial, governmental or military secrets, then the cryptographic strength must be substantial. However, maintaining such a high level of security and still keeping the real-time and limited-bandwidth constraints is not easy to accomplish.

Another form of lightweight encryption is so-called *scrambling*. Although scrambling is sometimes used to mean encryption, we clearly distinguish one from the other. By scrambling, we mean the filtered distortion of the analog output signal. At the receiving end, scrambled analog signal gets unscrambled using the inverse filter

transformation. Scrambling is usually achieved using an analog device to permute the signal in the time domain or to distort the signal in the frequency domain by applying filter banks or frequency converters. Although such transformations are in most cases not secure at all, analog signal scrambling is an interesting case study. Scrambling was simply a product of an immediate industrial need by cable companies for a fast solution to make the free viewing of paid cable channels more difficult than doing nothing at all. The scrambled signal possesses some entertainment value, however, not a great number of people purchased illegal unscrambling cable boxes that were fairly easy to manufacture. One of the reasons might probably be the lack of knowledge on behalf of the ordinary consumers, who did not realize how easy is to break a scrambled signal, or who thought that a cable company can somehow distinguish between those who use legal and those who use illegal unscrambling boxes. But most likely, people just didn't bother. When you really weigh the pros and cons, other than getting some free entertainment, there was no particular value gain. The cost of breaking the encryption system was simply higher than the value of encrypted information. In such cases, scrambling was enough to accomplish the desired effect. On the other hand, there are many multimedia applications where substantial security is of much higher priority. Obviously, using analog scrambling for a secret corporate video message from IBM to Microsoft is a bad idea. Using a conventional cryptosystem in combination with a selective encryption typically provides a much higher level of security.

So far, we have defined what we mean by degradation and scrambling. Both are simply types of lightweight encryption that usually provide a very low level of security. The main difference is that degradation distorts a digital multimedia file, while scrambling distorts an analog multimedia signal. This brings us to another point. In general, the success of scrambling may not guarantee the same success to the digital multimedia degradation. Let us not forget how much cheaper and how much more convenient it is these days to obtain software than it is to obtain hardware. If there is a freeware program that could break a particular degradation method, a consumer can very conveniently download it and use it to break the degraded multimedia. An Internet-based TV channel could be doomed of ever making consumers pay for viewing.

2.5 FORMAT COMPLIANCE

In many applications, it is desired that the encryption algorithm preserve the multimedia format. In other words, after encrypting the encoded multimedia, ordinary decoders can still decode it without crashing. This property of an encryption algorithm is often called *format compliance* (also called *transparency*, *transcodability* or *syntax-awareness*). When feeding the decoder with the format compliant encrypted data, the produced output seems distorted and randomized (see Figure 2.3). If the encryption algorithm was lightweight, the output often gives out some perceptual hints about the original content. However, if the encryption level was substantial, the output usually gives no perceptual information about the original multimedia.

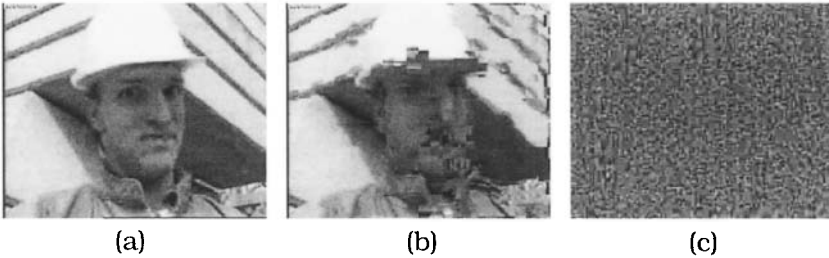


Figure 2.3. Decoded video produced by an ordinary decoder for (a) video without encryption, (b) video encrypted with lightweight format-compliant encryption, and (c) video encrypted with high-security format-compliant encryption.

Although format compliance is often desired property, for some applications it is not a critical one. In the next few chapters, we discuss both format compliant and non-format compliant multimedia encryption algorithms that are considered important.

2.6 CONSTANT BITRATE AND ERROR-TOLERANCE

In some applications, it is required that the encryption transformation preserves the size of a bitstream. This is known as the *constant bitrate* requirement. However, more often than not it is simply preferred that the output produced by an encryption-equipped encoder and the output produced by an ordinary encoder have similar sizes. That is, the encryption stage is allowed to slightly

increase the size of a bitstream. This is sometimes called a *near-constant bitrate*. A near-constant bitrate is likely to occur when block ciphers are used for encryption, since in that case the encrypted output is always a multiple of the blocksize.

For many multimedia systems *error-tolerance*, or *error-resilience*, is usually of high importance. Since the real-time transport of multimedia data often occurs in noisy environments, which is especially true in the case of wireless channels, the delivered media is prone to bit errors. If a cipher possesses a strong avalanche property, decryption will likely fail even if a single bit is flipped. The error-tolerance can be improved by applying some of the classical error-detecting or error-correcting codes. Unfortunately, these techniques are in many instances extremely costly to apply to an already bulky multimedia bitstream.

SUMMARY

Conventional cryptosystems are powerful but limited in terms of bulky and highly redundant multimedia data. Selective encryption and chaos-based encryption are often used to provide faster and better multimedia security. Various multimedia applications demand different encryption approaches. Consequently, choosing the appropriate security level is often not an easy task. Multimedia specific encryption algorithms are relatively new and the tools for their security assessment are not fully developed. Scrambling refers to a lightweight analog signal distortion, while degradation refers to a lightweight encryption of digital data where the encrypted signal is usually still perceptible. Format-compliance, error-tolerance and constant bitrate are important features of the multimedia encryption algorithms.



<http://www.springer.com/978-0-387-24425-9>

Multimedia Encryption and Watermarking

Furht, B.; Muharemagic, E.; Socek, D.

2005, XVII, 327 p., Hardcover

ISBN: 978-0-387-24425-9