NETWORK FRIENDLY MEDIA SECURITY: RATIONALES, SOLUTIONS, AND OPEN ISSUES[1]

Wenjun Zeng, Xinhua Zhuang, and Junqiang Lan
Dept. of Computer Science
Univ. of Missouri-Columbia
{zengw, jl629, zhuangx}@missouri.edu

ABSTRACT

Network friendly media security refers to the security technologies that are specifically designed to cope with existing and future multimedia networking infrastructures and technologies so as to ease the deployment and maintain or improve the quality of service performance of multimedia applications. It is especially useful for streaming and mobile multimedia applications where content adaptation is a necessity. In this paper, we analyze the various motivations behind network-friendly security solutions, review some of the most recent approaches, discuss some open issues, and suggest some potential solutions.

1. INTRODUCTION

Recent advances in networking and digital media technologies have created a large number of networked multimedia applications. Those applications and services are often deployed in a distributed network environment that makes multimedia contents vulnerable to piracy and malicious attacks. The security concerns, if not addressed appropriately, will potentially prevent or delay the wide dissemination of the multimedia applications. As a result, securing multimedia contents has been an active research area in recent years that involves such content protection technologies as encryption, authentication, and digital watermarking.

Multimedia data, unlike traditional data, exhibits several unique characteristics, including high data rate, power-hungry, time-constrained, synchronous, loss-tolerant, having components of different importance, and sometimes relatively low valued. Accordingly, advanced networking technologies have been designed specifically for the delivery of multimedia data. These unique properties of multimedia data and its distribution have posed significant challenges to conventional security technologies that were mainly designed for general data communication. To provide sufficient security at the same time reduce negative impact on the quality of service (QoS) performance and cost of the multimedia applications, especially streaming and mobile applications, media security technologies have to be carefully designed, preferably by explicitly considering the interplay between cryptographic, signal processing, and networking technologies. To that end, network friendly media security techniques have been recently proposed.

In this paper, we present various rationales behind network-friendly security solutions, review some of the most recent approaches, discuss some open issues, and suggest some potential solutions. The problem can also be partially viewed from the perspective of the placement of the security functions and its implications on the security and content adaptation functionality [1].

2. RATIONALES FOR NETWORK-FRIENDLY MEDIA SECURITY

There are several motivations behind media security solutions that provide network friendliness, as analyzed below.

Delay constraint

Traditional cryptographic algorithms such as DES and RSA, which were mainly designed for data security, are often not fast enough to process the vast amount of data generated by the multimedia applications to meet the real-time constraint. For example, to achieve secure interactive multi-point video communication, the huge amount of data created by multiple video streams, the real-time constraint by interactivity, and the complexity and structural characteristics of specific video coding technologies, together would require a special design and application of cryptographic algorithms. .

Power-constrained networks

In wireless multimedia communications, e.g., over advanced sensor networks, a high level security at extremely low computational cost will be crucial due to the battery-driven, low powered nature of the sensors and the handheld devices such as PDA and videophones.

Error-resiliency for lossy networks

Today's Internet and most wireless networks provide only the best-effort services, without certain QoS guarantees. That makes a secure transmission of real-time multimedia data even more challenging. Many standard cryptographic methodologies today require all bits to be received correctly. For example, any transmission bit error will render traditional authentication a failure. In addition, synchronization may become a problem for conventional security techniques in the case of packet loss [2]. This would imply a significant increase of latency due to the need of retransmission and/or the bit overhead caused by forward-error-correction. However, requiring all bits to be received correctly overlooks the fact that many video applications can tolerate certain bit errors or data loss that are perceptually less important. It is clear that traditional

[1] This research was supported in part by a grant from the University of Missouri System Research Board, and in part, by NIH grant DHHS 1 R01 DC04340-01A2 and NSF grant EIA 9911095.

cryptographic algorithms do not cope well with lossy networks and the loss-tolerant nature of the multimedia data.

Protocol friendliness

In a typical content delivery food chain, the multimedia is compressed and protected/encrypted at the encoder, and possibly stored in the disk, before it is sent to the server for delivery. Unfortunately, at the intermediate stages of such a delivery food chain, many technologies that are critical to the end-to-end quality delivery, such as transport packetization, unequal error protection, and random access, have been developed to work with unprotected compressed contents. For example, transport packetization using standard RTP (Real-time Transport Protocol) requires understanding the syntactical structures of the compressed bitstreams in order to assemble the RTP packets. The media security technologies should be designed to cope well with many existing and future network protocols.

Network adaptation and scalability

In many multimedia communication systems, to achieve certain level of end to end quality of service, application layer adaptation techniques such as congestion/rate control and media trans-coding performed by proxy servers located at gateways between heterogeneous networks (e.g., between the Internet and the cellular wireless networks [3]), or performed by the intermediate nodes in an overlay network [4], have been developed. Many of these existing techniques, e.g., transcoding, work on the unprotected bitstreams. Implementing security solutions at the intermediate nodes can address the network adaptation requirement, but would introduce processing overhead and may pose significant threats to the end-to-end security [5][6].

Content processing capability

In some applications such as multimedia content management and network traffic management, it would be desirable that content protection may still retain the feasibility of evaluating some statistical characteristics of the images such as global histogram, subband energy distribution, motion intensity information, etc., directly on the protected content [5].

Ease of deployment

Conventionally, cryptographic processing (e.g., decryption and then re-encryption, authenticity verification and then new authentication code generation) at the intermediate nodes is required for network adaptation techniques to function effectively. This will not only introduce significant processing overhead and delay at the intermediate nodes, but also impose significant challenges to the cryptographic system, especially the key management system, as the intermediate nodes will have to acquire the same cryptographic capability and to have access to the cryptographic keys. This will increase the deployment cost because there are a wide spectrum of cryptographic algorithms in use today, many of which are proprietary, and the intermediate network nodes would have to support many potential cryptographic algorithms [6].

End to end security implication

More seriously, requiring cryptographic processing at the intermediate nodes may compromise the end-to-end security [6][7]. This is because any inline decryption/re-encryption would surface clear content before it reaches the end-user. The conventional approach requires the intermediate nodes to be trusted and secure, which in some cases (e.g., in fully distributed peer to peer networks) may not be possible.

In summary, traditional cryptographic algorithms have not been designed to be friendly to multimedia networking and communication. The direct application of such security technologies may potentially render many of existing technologies in a content delivery food chain difficult to function. Clearly, there is a need for developing highly secure, lightweight, and network-friendly security solutions for quality delivery of compressed media streams. In particular, it is desirable to minimize the need to perform intermediate cryptographic processing on the protected content along the food chain of content delivery. One significant advantage of such approach is that many existing network infrastructures and standard techniques will be left intact, facilitating the deployment and adoption of the system. This would require that the content protection techniques be designed in a way such that various content adaptations at the intermediate stages of the food chain can be performed directly on the protected bitstreams, in transparency to the cryptographic operations..

3. A REVIEW OF EXISTING APPROACHES

There have been several approaches developed in recent years to address various network-friendly media security issues.

Selective encryption

To reduce the amount of processing overhead, selective encryption has been proposed. For example, in some selective encryption algorithms, only the intra-coded frames, i.e., the I frames, and Intra-coded Macroblocks (MB) of predictive coded frames are encrypted [8][9]. Selective encryption of sign bits of DCT coefficients and motion vector information has also been studied [10][5][6] . Previous work [11][5][6] has shown that the encryption of I frames alone or sign bits alone does not provide sufficient privacy/security. In fact, this is likely the case for any partial encryption schemes because of the information leakage from the unencrypted data. Nevertheless, these solutions are useful for applications that focus on introducing quality degradation rather than full secrecy.

Lightweight cryptographic techniques

Lightweight cryptographic tools have also been proposed that runs faster at possibly the expense of lower security. It is also possible to reduce the complexity without sacrificing the security, for example, by using XOR operation to reduce the rate of the data that is fed to the standard encryption function [12].

Joint encryption and compression

While most conventional approaches separate encryption from compression, a joint encryption with compression is possible that would generally lead to an improved system performance. The Zigzag-Permutation algorithm [13] is one of the earliest scrambling algorithms designed for joint MPEG video encryption/compression. The technique applies permutation to the DCT coefficients of the motion compensated residue image blocks so that they are encoded using an order different from the standard zigzag order. The algorithm, however, could result in a significantly lower compression ratio since the permutation destroys the statistics relied upon by a run-length coder. It also has security problems since the DC coefficients can be easily

recovered [12], and the spatial energy distribution of the image remains largely unchanged due to *local* permutation [5].

A joint encryption/compression framework was proposed in [5] where the transform coefficients are divided into blocks/segments and are subject to selective scrambling, which consists of all or some of the following operations: random sign change, block shuffling, block rotation, and coefficient shuffling within a subband segment, etc. The motion vectors, if any, are also subject to random sign change and shuffling. The scrambled coefficients and motion vectors are then subject to entropy coding and bitstream formation. Since coefficient blocks are spatially shuffled within a subband segment of the same frequency, the shuffling would not significantly change the statistics or degrade the coding efficiency, while the random change of the *high level* spatial configuration of *coefficient* blocks is difficult for the computer to restore without human intervention. The resultant encrypted bitstream conforms to the compression format, which provides some additional advantages as discussed below.

Format compliant selective encryption/scrambling

Recently we have also developed a format compliant selective encryption/scrambling framework that has been adopted by the MPEG-4 IPMP (Intellectual Property Management and Protection) Extension standard [14]. Format compliance is referred to as the property that the encrypted bitstream still looks like an unprotected compressed bitstream, and can often be played by the player (that renders unintelligent images). In this framework, encryption operations are executed after the entropy coding and full bit level compliance to the compressed video syntax can be maintained. One tool developed in this framework is to extract some fixed-length and variable length codewords, map them to *fixed length* index, and then encrypt the index using public encryption algorithms [6]. A primary concern of encrypting the variable length codewords using the aforementioned tool is the potential bit overhead. Another tool developed is to spatially shuffle codewords of the compressed bitstream in a way that the resultant bitstream would comply with the compression format as much as possible [2]. The tool requires the compressed bitstream be divided into groups of basic shuffling units, where each group will be shuffled, using a separate and dynamically updated shuffling table. Since the technique is simply a cryptographic key based re-organization of the compressed bitstream, it introduces no bit overhead. The security and error resiliency can be further enhanced using a self-synchronous dynamic shuffling table generation process [2].

The format compliant encryption framework would keep many carefully designed and desirable properties of the unprotected compressed bitstream, such as error resiliency, scalability, and protocol friendliness, unchanged. In addition, many random access, network bandwidth adaptation, and error control techniques that have been developed for unprotected bitstreams will still work with encrypted bitstreams, which is significant for multimedia delivery over time-varying lossy channels.

Scalable protection of scalable bitstreams

When the compressed media is created using scalable coding, scalable protection of the multimedia contents becomes feasible. Streams of different importance can be protected and possibly packetized separately so that less important packets can be simply dropped or truncated in case bandwidth is a problem [7]. One potential concern about such approach is that scalable video bitstreams are currently not widely supported in the industry yet.

Content-based image authentication

Traditional cryptographic authentication schemes using digital signatures or message authentication codes are designed to ensure that no single bit alteration is allowed. However, in many multimedia applications, it is desirable to tolerate some content preserving operations such as compression, while being capable of detecting other malicious attacks. An authentication system that distinguishes content preserving operations from malicious modifications is referred to as a content integrity system [15][16]. Among others, semi-fragile digital watermark has been proposed as one of the tools for such purpose. A semi-fragile watermark can be used to detect and locate the modifications to the host media [15]. When the host media is tampered, the watermarks will also be modified that can serve to detect the tamper or even locate where exactly the host media is tampered. Previous work, however, has been focused on still image tamper detection, and has mainly addressed the survivability of the semi-fragile watermark to still image compression [15].

Multicast authentication, key management and watermarking

Multicast of multimedia data is a bandwidth efficient delivery mechanism. Both IP multicast and application layer multicast have been studied. Multicast security poses a number of challenges because the shared group key has to be updated frequently, e.g., whenever a member joins/leaves the group. Complexity, scalability, loss resilience, and synchronization are some of the critical issues to be addressed for key distribution. Using distributed watermarking for finger printing individual receivers has also been studied. A review of multimedia security in group communications can be found in [17].

4. SOME OPEN ISSUES AND DIRECTIONS

Despite recent progress in network-friendly security technologies, there are still many open issues regarding the complexity, security, quality of service performance, and ease of deployment. For example, synchronization between key stream and media stream is a common problem for delivery over lossy networks. In the following, we identify a few of them and suggest some potential solutions.

Semi- format compliant encryption

As discussed above, fully format compliant encryption schemes are most friendly to existing modules in an end-to-end communication system. However, a fully format compliant solution could be over-restrictive for some applications, thus sacrificing some performance (e.g., bit rate overhead and/or security level). In fact, in many applications, the preservation of *some* important syntactical structures of the compression format that are required for network processing would be sufficient. We refer to this property as *semi-format compliance*. For example, for RTP packetization of MPEG-4 video, only syntatical-structures above the video packet level are required. Accordingly, it might be more efficient to devise some compromised solutions where only some most wanted properties (e.g., transport protocol compatibility, scalability, error resiliency) are preserved [2][7][18].

It remains to find out the best security architecture for semi-format compliant encryption such that more important multimedia data will be delivered in a more efficient, reliable, and secure way, especially for non- fully scalable compression formats that are currently more widely adopted.

Multimedia content authentication/integrity

In a hostile environment, the complication introduced by potential data loss due to bandwidth adaptation and/or transmission errors requires more intelligent tools for integrity checks and authentication of multimedia data. These tools need to be robust to transmission loss while being sensitive to intentional modifications/attacks. One way to address this issue is to exploit the unique properties of the watermark-based authentication scheme. This approach has the advantage and capability of performing "local" authentication at the application layer without necessarily binding a large amount of data together. It is extremely useful in the case of packet loss or random bit error, where it can still provide authentication to the correctly received data without being affected by the lost data. A further advantage is that, unlike asymmetric encryption based authentication, it maintains full format-compliance of the media stream, which provides the greatest flexibility for adaptation of the protected content [1].

In particular, the followings are a few interesting and challenging problems that warrant further investigation: 1) How to redefine the metric for content authenticity given that data loss is tolerable and acceptable? 2) How to redefine and fully address the integrity of multimedia/video content in a lossy network setting? This appears to be an area that has not been well studied. For example, one particular challenging problem is how to differentiate packet loss due to network congestion or random bit errors from packet loss due to intentional deletion. We outline a potential solution here. We believe that in a lossy environment, it is necessary to require a (application dependent) minimum amount of content information be correctly received and authenticated. For example, it may correspond to a visual summarization of the video content using the key frames. It may be critical to verify the integrity of the extracted video scene structures and their ordering and temporal-dependency, which may require error-free delivery and full protection of such important information using an appropriately designed protocol.

Authenticated multicast of real-time multimedia

In a multicast setting, typically asymmetric mechanisms are required for authentication, as the receivers may be mutually un-trusted, which renders a shared group key inapplicable. However, traditional asymmetric authentication schemes are too slow for use in multimedia distribution, especially for low powered networks and devices. This poses a significant challenge. Although for non-interactive applications, asymmetric authenticated broadcast has been constructed from symmetric primitives using delayed key disclosure and one-way function key chains [19], it remains an open issue for real-time interactive multimedia applications where delay is extremely critical.

We conclude that, to achieve optimized secure multimedia communication, it is crucial to take full advantages of the interplay between cryptographic, signal processing and networking technologies. The effects of the underlying network characteristics and the implication of existing network infrastructures should be taken into account in designing secure multimedia communication algorithms.

REFERENCES

[1] W. Zeng, J. Lan and X. Zhuang, "Architectures and analysis for content protection of adapted media", submitted to *IEEE Trans. CSVT*, June 2004.

[2] W. Zeng, J. Wen and M. Severa, "Fast self-synchronous content scrambling by spatially shuffling codewords of compressed bitstreams," *Proc. IEEE ICIP.*, Sept. 2002.

[3] G. Cheung, W. Tan, and T. Yoshimura, "Double feedback streaming agent for real-time delivery of media over 3G wireless networks," to appear in *IEEE Trans. Multimedia*, *Special Issue on Streaming Media*, April, 2004.

[4] Y. Chu, S. Rao, and H. Zhang, "A case for end system multicast," *Proc. ACM SIGMETRICS Conference*, June 20.

[5] W. Zeng and S. Lei, "Efficient frequency domain selective scrambling of digital video", *IEEE Tran. Multimedia,* vol. 5, no. 1, pp. 118-129, March 2003. A preliminary version also in *Proc. ACM Multimedia* , Nov. 1999.

[6] J. Wen, M. Severa, W. Zeng, M. Luttrell and W. Jin, "A format compliant configurable encryption framework for access control of video," *IEEE Tran. Cir. & Sys. for Video Tech., Special Issue on Wireless Video,* pp. 545-557, June 2002. A prelim. version in *IEEE Workshop MMSP*, 2001.

[7] S. Wee and J. Apostolopoulos, "Secure scalable streaming enabling transcoding without decryption," *IEEE International Conf. Image Proc.*, vol. 1, pp. 437-440, 2001.

[8] T. Maples and G. Spanos, "Performance study of a selective encryption scheme for the security of networked, real-time video," *Proc. 4th Inter. Conf. Computer Communications and Networks*, Las Vegas, Nevada, Sept. 1995.

[9] J. Meyer and F. Gadegast, "Security mechanisms for multimedia data with the example MPEG-1 video," http://www.cs.tuberlin.de/phade/phade/secmpeg.html, 1995.

[10]C. Shi and B. Bhargava, "A fast MPEG video encryption algorithm," *Proc. ACM Multimedia*, pp. 81-88, 1998.

[11]I. Agi and L. Gong, "An empirical study of secure MPEG video transmissions," *The Internet Society Symposium on Network and Distributed System Security*, Feb. 1996.

[12]L. Qiao and K. Nahrstedt, "Comparison of MPEG encryption algorithms," *Inter. J. on Computer & Graphics,* 22(3), 1998.

[13]L. Tang, "Methods for encrypting and decrypting MPEG video data efficiently," *Proc. ACM Multimedia*, 1996.

[14]MPEG4 IPMP FPDAM, ISO/IEC 14496-1:2001/ AMD3, *ISO/IEC JTC 1/SC 29/WG11 N4701*, March 2002.

[15]C. Lin and S. Chang, "Semi fragile watermarking for authentication of JPEG visual content," *Proc. SPIE Inter. Conf. Security and Watermarking of Multimedia Contents*, vol. 3971, Jan. 2000.

[16]M. P. Queluz, "Towards robust, content based techniques for image authentication," *IEEE Workshop MMSP,* 1998.

[17]A. M. Eskicioglu, "Multimedia security in group communications: recent progress in key management, authentication, and watermarking," *Multimedia Systems* 9:239-248, Springer-Verlag, 2003.

[18]M. Wu and Y. Mao, "Communication-friendly encryption of multimedia," *IEEE Workshop MMSP*, 2002.

[19]A. Perrig, et al., "SPINS: Security protocols for sensor networks," *ACM Mobile Computing and Networking,* 2001.