Chapter 7
Challenges and Solutions in Multimedia Document Authentication

Stefan Katzenbeisser
Security Engineering Group, Technische Universität, Germany

Huajian Liu
Fraunhofer-Institute for Secure Information Technology, Germany

Martin Steinebach
Fraunhofer-Institute for Secure Information Technology, Germany

ABSTRACT

Multimedia document authentication allows the judging of the authenticity and integrity of digital documents. Today a variety of such tools exist which are constructed using different approaches, namely forensic methods, perceptual hashes and digital watermarks. Digital document forensics identifies the source of a document as well as its integrity by means of passive estimation. Perceptual hashing allows computing of short digests of documents; the hashes are insensitive against some signal processing operations and may serve as proof of integrity. Finally, authentication watermarking even allows gaining precise and reliable knowledge on the nature of modifications that a digital document underwent. In this chapter, we give an overview of the three complementary technologies, survey state-of-the-art methods and provide an analysis of their strength and weaknesses.

MOTIVATION

Multimedia data becomes more and more relevant for applications that require a certain level of trust in the integrity and the authenticity of documents. Examples include scanned contracts and documents which integrity needs to be verified, photos or video clips attached to news reports which contents should be provably authentic or recordings of interviews which shall be used as evidence in the future. The possibility to store documents in digital form raises new challenges with respect to the recognition and

DOI: 10.4018/978-1-60566-836-9.ch007

prevention of forgeries and manipulations. By using a powerful personal computer and sophisticated image editing software, even an inexperienced user is able to edit a picture at will, e.g. by adding, deleting or replacing specific objects, thereby creating "perfect" manipulations that do not introduce visually noticeable traces (Cox et al., 2001; Zhu et al., 2004). It is very hard, if not impossible, for a human to judge whether a multimedia document is authentic only by visual inspection. As a result, the old proverb "words are but wind, but seeing is believing" is not true anymore in the digital era.

Multimedia document authentication tries to alleviate this problem by providing tools that verify the integrity and authenticity of multimedia files. In particular those tools detect whether a document has undergone any tampering since it has been created (Zhu et al., 2004). In this chapter we focus on tools that operate on raw data (such as sequences of image pixels or audio samples) instead of compound multimedia objects, as this is the focus of current research. Depending on the application scenario, three different approaches – media forensics, perceptual hashing and digital watermarking – can be found in the literature.

The field of *media forensics* tries to examine a multimedia document in order to decide whether it is authentic or not. No prior knowledge on the document is assumed. Technically, these schemes look for suspicious patterns that indicate specific tampering. In addition, it is sometimes possible to determine the device that was used to create the document (such as a scanner or camera). Note that document forensics differs fundamentally from steganalysis. The latter tries to detect and decode any secret imperceptible messages encoded within a document, while forensics deals with the examination of document authenticity and integrity; steganalysis is thus out of scope of this chapter.

While promising approaches exist to uncover tampering, more reliable results can be achieved if a potentially tampered document can be compared to its "original" version. This operation is usually harder than it seems, as media documents may undergo several processing steps during their lifetime; while these operations do not modify the visual content of a document, its binary representation does change. For example, media files are usually stored and distributed in compressed form. Such compression methods are often lossy and will render the decompressed data slightly different from the original copy (e.g. the JPEG format does not store perceptually insignificant parts of an image). Besides compression, the data may also undergo other incidental distortions such as scaling. Thus, the binary representation of media documents cannot directly be compared to each other. *Perceptual hashes* provide an automated way of deciding whether two media files are still "perceptually identical", for example whether one document is a copy of another one, which was processed without changing its semantics. A hash is a short digest of a message, which is sensitive to modifications: if a document is severely changed, the hash value will change in a random manner. Hashes can be used to verify the integrity of an object if the hash of the "original" is stored at a trustworthy place, such as a notary. During verification, the document is hashed and the hash is compared to the hash of the original. If the hash differs, the document is assumed to be modified.

In cryptography, several hash functions have been proposed. However, they are usually unsuited to the authentication of media files, as they provide only bitwise authentication. The targeted data must be identical to the original copy in order to be considered as authentic; even one bit difference will render the whole content unauthentic. As mentioned above, this is inappropriate for media files. A conventional cryptographic hash function thus cannot distinguish between incidental and intentional distortions due to malicious manipulations of the content. In contrast to cryptographic hashes, perceptual hashes allow to compute a digest of a document that remains invariant under some distortions that do not alter the semantics of the document. Thus, processed documents can still be reliably compared to each other. The