Evaluating trusted electronic documents

Petr Švéda

xsveda@fi.muni.cz

Department of Program Systems and Communications
Faculty of Informatics MU Brno
Botanická 68a, 602 00 Brno, Czech Republic

Abstract

An attack does not have to be the biggest threat to a digital signature system or application. The real threat is failing, which implements the security required by law to establish trust. Today security and reliability have become the key for digital signature and electronic documents. In essence, a technical problem has become also a legal issue.

Proprietary solutions are not compatible and their security depends on a closed design. Real solutions based on XML standards are solving only the problem of a content verification partially. It is one of the clue points, but other unsolved or partially solved problems are left there – e.g. content presentation, trust issues in signature creation, long-term signature verification, time stamps, compatibility and legal issues. A lot of drafts and partially pre-published standards exist there, which solve only several isolated problems. Any complete evaluation concept has not been published yet. This paper presents the evaluation approach based on requirements on a content, context and structure of a signed electronic document.

Keywords: content, context, digital signature, electronic document, evaluation, structure, trust.

1 Introduction

People trust data in the context of a document. It is possible to secure an electronic document by digital signature techniques – this document would be trusted and unalterable. The key points of an electronic document are flexibility and editability. So it is infeasible to trust an editable file format, which has a lot of optional settings that can change a visual representation of an electronic document dramatically.

Existing file and data formats can be divided (according to the data structure) into three groups:

- *Mark-up based* – Formats that capture logical structure and may include some necessary metadata for viewable transformation (for instance XML or TeX file formats).
- *Page describing oriented* – Formats that capture layout, e.g. Portable Document Format (PDF) or PostScript (PS).
- *Combined* – Formats that contain mixture of document's layout and structure. An example is Rich Text Format (RTF) or Microsoft Word Document (DOC).

A fundamental conflict exists between trust and usability in current combined file and data formats. Mark-up based formats can be trusted if the problem of an unambiguous presentation is solved correctly. Detailed discussion about trust issues related to XML documents can be found in [7].

It seems to be suitable to separate content data and its presentation for a trustworthy document structure. It has to contain a signed instance of an electronic document, which was visible on a signer's screen at the moment of signing. A processing viewer or editor uses the trusted signed instance, called view, later. Design principles for a trustworthy document structure, which is based on views, are proposed in [11].

2 Trustworthy documents

It is necessary to preserve the *content*, *context* and *structure* of a document to remain reliable and authentic. A trustworthy document preserves the actual content. There is also is required information about the document that relates to the context in which it was created and used. Specific contextual information will vary according to

the requirements of the activity. It is also necessary to preserve the structure or arrangement of the document. The failure in the structure of the document will impair its structural integrity. That, in turn, may undermine the reliability and authenticity of the document.

There are special considerations when dealing with the preservation of the content, context, and structure of documents that are augmented by digital signatures:

- *Content* – The digital signature or signatures in a document are part of the content. They indicate who signed a document. Sometimes they also can show that person approved the content of the document. Multiple signatures can indicate initial approval and subsequent concurrences. Signatures are often accompanied by some other information (e.g., creation date). All these things are part of the content of the document and needs to be preserved. The lack of this information seriously affects a reliability and authenticity of the document.
- *Context* – Digital signature technologies rely on individual identifiers that are not embedded in the content of the document. Trust paths, time stamps, and other means are included there to create and verify the validity of a signature (see Section 3). This information is outside of the content of the document, nevertheless it is important to the context. It provides additional evidence to support the reliability and authenticity of the document. The lack of these contextual elements affects seriously one's ability to verify the validity of the signed content.
- *Structure* – Preserving the structure of a document means its physical and logical format. This is also the question of the relationships between the data elements comprising the document which remains intact physically and logically. It is essential that all states of the document structure are fully and unforgeably described. It is necessary to assure the integrity of all information about structure of the document.

3 Revalidation issues

The information necessary for revalidation (i.e., the public key used to validate the signature, the certificate related to that key, and the certificate revocation list from the certificate authority that corresponds to the time of signing) must be retained for as long as the digitally-signed document is retained. Both contextual and structural information of the document must be retained.

Important contextual information are:

- *Certificate (or public key certificate)* – A digitally signed data structure that binds the identity of a certificate holder to a public key. It is defined in the X.509 standard [6].
- *Certificate policy* – A named set of rules that indicates the applicability of a certificate to a particular community and/or class of an application with common security requirements. For example, a particular certificate policy might indicate applicability of a type of certificate to the authentication of electronic data interchange transactions for the trading of goods within a given price range.
- *Certificate practice statements* – A statement of the practices which a certification authority employs in issuing certificates. It provides a detailed explanation how the certificate authority manages the certificates, its issues and associate services, such as key management. The CPS acts as a contact between the certification authority and users, describing the obligations and legal limitations and setting the foundation for future audits.
- *Certificate revocation list* – A list of revoked but unexpired certificates issued by a certification authority. A list of subscribers paired with their digital signature status and the reason for the revocation.
- *Trust paths* – A chain of certificates of trusted third parties among parties to a transaction which ends with the issuance of a certificate that the relying party trusts.
- *Trust verification records* – Records that prove when and how the authenticity of the signature was verified. An example of this would be an Online Certificate Status Protocol (OCSP) [9] or another response from a certification authority.

Structural information are all the cryptographic primitives, file and signature format properties (e.g., necessary metadata tags) [2, 4, 5]. A digital signature remains valid as long as all the cryptographic primitives (e.g., hash functions, encryption algorithm and digital signature schemes) and parameters (e.g., key material and

certificates) remain valid. If one of these components becomes invalid, then the signature would lose its property as evidence. As a consequence, the signature verification process cannot succeed.

3.1 Procedures of signature validation

Consider the basic digital signature scheme with an appendix (e.g., DSA, ElGamal, Schnorr). There also exist digital signature schemes with message recovery (e.g., RSA, Rabin, Nyberg-Rueppel). The later type can be exchanged for the former one [8, 10]. The recipient uses an appropriate sender's public key to decrypt the attached signature, computes the hash value of a received message and compares both characteristics. If they are equal, the signature (and document) is verified. Aside from the basic scheme, there are distinguished other validation schemes in [1]:

- *Initial signature verification* – It is the action of capturing the information that makes the digital signature verifiable against a signature policy. This should be done "soon after" a digital signature is generated.
- *Usual signature verification* – It is the action of checking a digital signature against a signature policy. This may be done at any time after the initial signature verification (e.g. years after the digital signature was produced).
- *Archival signature verification* – It is the action of checking a digital signature against the information that were secure and valid at the time of the signature, but which are likely to be no longer secure at the time of a later verification.

4 Evaluation concept

This proposal of the evaluation divides signed documents into four categories according to the level of the trustworthiness they can offer. The higher category the higher assurance of reliability and authenticity of signed documents they offer. The choice of level depends on the type of an application, potential threats and on the security functional requirements.

- *Level 0* – Documents falling into this category are more or less simple. They only offer very restricted level of reliability and authenticity. And they can be "easily" cheated or impeached. Trustworthiness of an electronic document is based only on a reliable public key issuing. A typical example of such document can be a plaintext file or email signed by popular PGP package or with a certificate issued only on e-mail address.
- *Level 1* – Electronic documents at level two require signature created with certificate. The requirement on an enforced proof of the origin can be supplied by a qualified certificate [3]. The verification is sufficient initial signature verification. A structure has to contain also a document snapshot, which was visible on the screen at the moment of signing. An example of such document can be an MS Word or Excel document signed by a commercial solution that stores virtual printed copy. Tenths of possible product solutions are available.
- *Level 2* – Level two documents require structure that allows verification of content data. Exposed components of the system which are working with a document (typically editor, viewer and signature creation or verification application) have to do integrity checking. The verification is sufficient usual signature verification. A document has ability to contain audit data (e.g., revisions). An example of such document can be a XML document signed by XAdES format [4].
- *Level 3* – Archival signature verification is required for the documents on level three. The document structure allows content verification by two or more techniques. A supplemental public interface to the document structure must be available. It is necessary to monitor all data integrity. All previous states of the document are described fully and unforgeably. The electronic document falling into the level three should be able to resist even well-funded attacks. But it is rather difficult to design a document resistant to conspiracy among other parties inclusive TTP against the last one.

	Content requirements	Context requirements	Structure requirements	Other requirements
Level 0	no requirements	basic verification	no requirements	public key issuing
Level 1	basic user identity generation	initial signature verification	contains document snapshot created during signature creation process	enforced proof of origin and its identity
Level 2	audit data generation	usual signature verification	allows content verification	all data integrity checking
Level 3	full audit data availability	archival signature verification	allows content verification by two or more techniques with public interface	all data integrity monitoring

Table 1: Brief overview of evaluation requirements.

5 Conclusions

The law, regulations and standards do not clearly identify which technology has to be used to implement digital signatures nowadays. Digital signatures techniques, known from the field of cryptography, assure legal signature requirements on the binary data level. Human beings cannot read binary data and do not understand them. People depend on widely spread file and data formats.

Widely spread combined file and data formats cannot be trusted. So the trustworthiness criteria have to be set. This paper presents the evaluation approach based on requirements on a content, context and structure of a signed electronic document. The proposed evaluation approach can be extended via + notation. For an example, level one+ document can correspond to some but not all requirements for level two documents.

References

[1] CEN: Procedures for Electronic Signature Verification, CWA 14171, 2001.

[2] ETSI: Electronic Signature Formats, TS 101 733, 2000.

[3] ETSI: Qualified certificate profile, TS 101 862, 2001.

[4] ETSI: XML Advanced Electronic Signatures (XAdES), TS 101 903, 2002.

[5] ETSI: XML Format for Signature Policies, TR 102 038, 2002.

[6] ITU: X.509, The Directory: Public-key and attribute certificate frameworks, 2000.

[7] N. Lundblad: Trusted Documents, in XML Europe 2001, pp. 27-34, 2001.

[8] A. J. Menezes, P. C. Oorschot, and S. A. Vanstone: Handbook of Applied Cryptography, CRC Press, ISBN 0-8493-8523-7, 1996.

[9] M. Myers, R. Ankney, A. Malpani, S. Galperin, and C. Adams: X.509 Internet Public Key Infrastructure Online Certificate Status Protocol, RFC 2560, 1999.

[10] B. Schneier: Applied Cryptography, John Wiley & Sons, Inc., ISBN 0-471-11709-9, 1996.

[11] P. Švéda: Trustworthiness of Signed Data, Technical report, Faculty of Informatics, Masaryk University, 2002.