Practical considerations in securing electronic signatures

NICK POPE

This paper takes a general look at the practical requirements for securing electronic signatures in both the physical as well as the electronic domains. It suggests that, whilst digital signatures have an important role to play in ensuring the security within the electronic domain, consideration also needs to be given to security in the physical world. Also, it is suggested that security applied to electronic signatures should be cost effective with a balanced approach to all the risks.

Background

Ever since their discovery, public key techniques and digital signatures have been considered as one of the prime means of providing the electronic equivalent to a physical signature on a piece of paper. Early United States state legislation gave specific recognition to the use of digital signatures. However, latterly the United States has moved to a more general approach, as illustrated in the US Federal Electronic Signatures in Global and National Commerce Act, 15 U.S.C. §§ 7001-7003, in which the term electronic signature is defined, in section 106 (5) as: "an electronic sound, symbol, or process, attached to or logically associated with a contract or other record and executed or adopted by a person with the intent to sign the record."

The European Union Directive,[1] and European legislation following the Directive has a dual approach. First, it recognises a specific "qualified" form of electronic signature that in its practical realisation is based upon the use of digital signatures, supported by a "qualified" certificate, and a smart card device use to hold the signing key. The second alternative, referred to in article 5.2 of the EU Directive, is neutral in the form an electronic signature can take, based on the definition of the general term electronic signature as provided for in article 2(1) "data in electronic form which are attached to or logically associated with other electronic data and which serve as a method of authentication".

In some European countries where there has been significant government support for the adoption of digital signature technology, there has been some acceptance of the use of smart cards with digital signatures.[2] However, the use of digital signatures has yet to have a significant effect on the general on-line electronic commerce market, where the additional costs associated with the use of digital signatures and smart cards have to show clear benefits before being taken up. Often retailers will use a simple 'click' button supported by some private information such as mother's maiden name as sufficient to indicate agreement to some legal conditions or a contract. The use of such very basic authentication techniques is clearly open to abuse. Once such private information is given to a site that is later found to be fraudulent or operates lax security, it can no longer be considered private. This concern with minimising costs is leaving users open to significant risks.

At the moment there is no mid-point between digital signature technology that can be expensive to deploy and operate, and leaving users with only very basic protection. It is suggested that there is a need for a means of applying signatures that are cheaper to deploy and operate than the use of digital signatures supported by smart cards, but has a greater degree of security than the very basic "click to agree" mechanisms.

In some European countries where there has been significant government support for the adoption of digital signature technology, there has been some acceptance of the use of smart cards with digital signatures

[1] Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures (OJ 19.1.2000 L13/12).

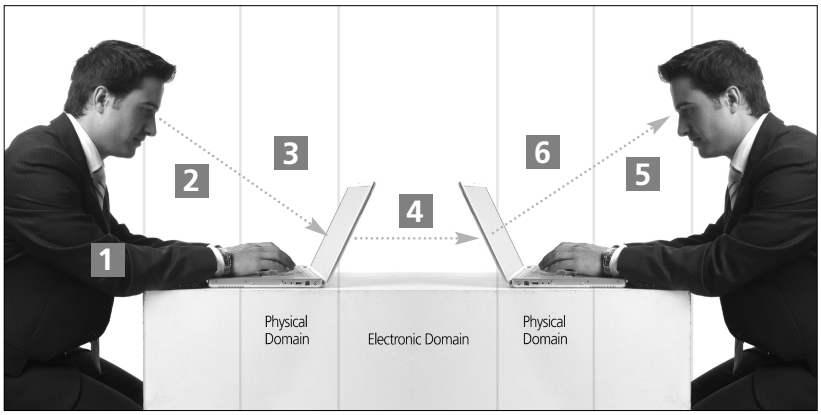
[2] For instance, see Ing. Franco Ruggieri, 'A technician's views on the digital signature in Italy', *e-Signature Law Journal*, 2005, Volume 2 Number 1, 53 – 59.

Whilst biometric features cannot be easily copied in the physical domain, once converted to the electronic domain they are just a string of numbers that can be easily replicated

Considering the physical as well as electronic domain

Current considerations regarding the requirements for security of electronic signatures, such as specified in the EU Directive, gives detailed regard to the signature in the electronic domain, but little thought to the physical world. The authentication provided by techniques such as digital signatures needs to be related to physical people. Signing keys are only an abstract sequence of bits, and they need to be securely linked to the physical person who owns the key. Similarly, digitally signed electronic documents are also only an abstract sequence of bits, they in turn need to be converted to something directly visible that a human can understand. The action of signing needs to be under the control of the signatory so that it can be used to indicate the intention of the signatory.

In considering the security of an electronic signature in a holistic way, it is necessary to address a range of issues as identified below:



1 Authentication of the signatory

In order for an electronic signature to be shown to belong to a specific individual, it is necessary to authenticate the signatory. There are three aspects of a human that are commonly used for authentication:

- **Something I have:** A smart card or some token, a computer or smart telephone or some other object owned by the person being authenticated.
- **Something I know:** A password, PIN code or some secret known only to the person being authenticated. This can include basic knowledge such as a mother's maiden name.
- **Something I am:** Biometric features such as fingerprints, facial features or retina scan.

Each of these classes of authentication has their own drawbacks. Something I have can be stolen and used by someone else. Something I know can be revealed to someone else. Whilst biometric features cannot be easily copied in the physical domain, once converted to the electronic domain they are just a string of numbers that can be easily replicated.

So-called 'strong' methods for authenticating users often combine at least two of the above. For example, smart cards (something I have) are generally only acceptable if provided with a PIN (something I know). However, it should be noted that such strong techniques could have significant cost implications over simpler techniques such as authentication based on just something I know (for example, passwords). Also, it should be noted that when used for digital signatures, although the private signing key in the card may be what is used to sign electronically, it is really the physical device and the PIN that authenticate the user.

2 What the signatory sees is what they sign

When signing a document, signatories do not sign the document blind. They need to see the document to know that it is acceptable before applying a signature to indicate, for example, their agreement. The document stored electronically and sent to the recipient generally should properly represent the information as seen by the user. This should include the formatting of the information. Thus the document should be stored in a form that properly represents the formatted document as seen.

As an aside, the term 'viewed' can be taken as audible as well than a visible representation of a document. Where a viewer converts a document to spoken words, this may be considered as viewing the document in audible form. In such cases, it may be necessary to record that signed document was 'viewed' in audible form to avoid any ambiguity over what is being signed.

This requirement implies first, that the method of encoding the data should include explicit formatting so there can be no issue over what is seen. The Adobe Acrobat PDF format is a good example of an encoding technique that unambiguously defines the format and layout of

the data. This is less so with Word that depends to an extent of the layout of the users printer. The HTML web format language is also generally good, although the layout can vary due to the local browser settings. However, the XML data encoding technique on its own does not include formatting, and so requires the addition of a 'style-sheet' when signed to ensure that the property of "what is seen is what is signed" is preserved in the encoding of the data.

Another issue that needs to be considered is the use of scripts within the document. Such scripts significantly increase the risk of the encoded document being viewed differently if different situations. Most modern encoding techniques (PDF, HTML, Word) include scripting facilities. Wherever possible, signed documents should avoid the inclusion of any scripts, and whenever signed documents are being viewed which contain scripts, the application should give the user a warning that, because of the use of scripts, it may not be guaranteed that the document is viewed as originally signed.

Finally, there needs to be some assurance that the software on the users personal computer is operating correctly and that there is no rogue or malicious software that affects the document display. This is discussed later in this paper.

3 Indication of intent

Whilst intent is not explicitly identified as a requirement in the EU Directive, unlike its US counterpart, if a signature is to be used to indicate agreement, there needs to be some control over the act of signature creation. This could be through the clicking of an 'I agree' button or some other user input into the application which applies the signature, or through the same means as used for authentication. When using a smart card and PIN, the entry of the PIN can be used to indicate intent.

4 Integrity protection within the electronic environment

For a signed document to be sent and held by other parties for later use as proof of agreement, the integrity of the signature with the document needs to be maintained. If either party disputes the document signature or its content, some means of proof that the integrity of the document has been preserved must be shown.

Digital signatures are probably the best mechanism for doing this, although not the only one. The significant advantage of digital signatures is that only one element of the signing environment needs to be kept absolutely secure: the private key. Once protected with a digital signature, the data can be distributed and copied, and any tampering is immediately evident. The main impediment to the use of digital signatures appears to be the costs and management overhead involved in putting the infrastructure into place to provide keys securely to those who wish to sign.

Other techniques exist which can be used to provide proof of the integrity of data in the electronic environment. This commonly involves a trusted third party to provide the necessary integrity protection. This can include use of time-stamping, electronic notary services and electronic seals.

5 Viewing and validating

Any party relying on the document or providing adjudication on the validity of the document needs to have some assurance that the document being viewed is an authentic signed document. This requires assurance that:

- There is some visible indication that the protection applied in the electronic domain confirms integrity and authenticity of the document after it has been signed.
- There is some means of tracing back to some physical authentication of the signatory and the 'intent to sign' bound in some way to the electronic mechanism. For example, there is a digital certificate which links the public signing key to a person who purports to have sole control over the public key. However, again it should be recognised that this is not the only means of authenticating the signatory. An alternative approach would be to bind some means of authentication, a biometric information or one-time password, into the document.
- The document is displayed to the relying party or adjudicator in a way that conveys the same information as shown to the signatory.

6 Personal computer security

Perhaps the greatest security vulnerability lies in the platform used by the signatory and the party

Wherever possible, signed documents should avoid the inclusion of any scripts, and whenever signed documents are being viewed which contain scripts, the application should give the user warning that, because of the use of scripts, it may not be guaranteed that the document is viewed as originally signed

viewing and relying on the signed document. Even if high security smart devices and keys are used for digital signatures, if the security of the environment that displays the document and applies the signature is uncertain, its validity can be open to dispute. Even in commercial environments with good security practices, there have been occurrences of malicious software such as viruses that have widely infected their IT systems.

Thinking outside the conventional digital signature box

Given that even with the use of high security cards and strong signature algorithms, there remains significant risk, a more cost effective solution is worthy of attention. There exist a range of alternative technologies that warrant consideration.

■ Biometrics

Biometrics have the considerable advantage of being strongly linked to an individual person. Fingerprints and such like are very good for authenticating an individual in the physical domain. However, once biometric features are in an electronic form, they are like any piece of electronic information, subject to misuse. Once in an electronic form, they need protection to bind them securely to the signed document. Digital signatures can have a useful role to play here, but this form of signature need not necessary belong to an individual, but rather a trusted device which binds the biometrics as a signature on behalf of the user.

■ One-time passwords

One-time passwords, as the name implies, can only be used once. For example, they can be created automatically from a special purpose hand held device, or for infrequent use for special purposes provided in printed form. Coutts & Co the bankers in the United Kingdom use a mixture of one-time passwords with two-factor authentication, and JP Morgan Chase are working with RSA security on the use of one-time passwords for signing. A one-time password has a significant advantage over simple passwords, in that once a password has been used for a specific purpose, it cannot be re-used. Whilst the threats are less in the electronic domain than with biometric solutions, the use of third party signatures or time-stamps would minimise the threats of misuse.

■ Use of existing personal devices

Smart telephones or other personal devices, whilst not being as secure as special purpose smart card devices, have the advantage as already being 'something I have' which users are likely to take great care that the device remains in their personal possession.

■ Third party digital signing services

Devices that apply signatures on behalf of users can be a cost effective way of ensuring that data is protected whilst in the electronic domain. A range of such devices exists, including time-stamping servers, digital notaries, and electronic seals. Whilst they do not directly authenticate the source, they can be used to bind authentication data (such as one-time passwords) to the document, and in some cases provide indirect authentication.

■ Signature gateway

A signature gateway provides a conduit between special purpose forms of signature (for instance, based on biometrics) to more widely recognised form of signature. For example, such a gateway may take a biometric signature applied locally, and counter-sign the document on behalf of the original signatory, using a standard form of digital signature.

Conclusions

A number of potential alternative mechanisms for providing electronic signatures have been outlined in the discussions above. The use of digital signatures is becoming well established where the support is available to set up the necessary infrastructure. However, consideration should be given to the use of alternative mechanisms that are more cost effective. In addition, it should be recognised that even with the use of digital signature and smart card technologies, there can remain a number of security issues to be addressed. The major risks that may remain on any unsecured platform which is used to view and sign documents, may negate any benefits gained from using secure signature mechanisms instead of much simpler and cheaper techniques. ■

Nick Pope is an independent consultant in IT security who has been working on electronic signatures over a number of years. He is co-chair of the OASIS Digital Signature Services technical committee, lead expert on European Certificate Policy standardisation and a member of the editorial board of the Journal.

pope@secstan.com