REVIEW OF THE ELECTRONIC TRANSACTIONS ORDINANCE: CAN THE PERSONAL IDENTIFICATION NUMBER REPLACE THE DIGITAL SIGNATURE?

*K. H. Pun, Lucas Hui, K. P. Chow, W. W. Tsang, C. F. Chong and H. W. Chan**

In a recent consultation document, the Information Technology and Broadcasting Bureau proposed that personal identification numbers (PINs) be accepted as a form of signature for the purposes of the Electronic Transactions Ordinance (ETO) (Cap 553). This article explains why this proposal is fundamentally flawed. The article identifies three basic requirements for a signature and examines whether they are satisfied by digital signatures and PINs. It concludes that while a digital signature has built into it all the elements necessary for compliance with the requirements, a PIN can only be used for the purpose of authorisation and cannot be elevated to the status of a signature as required by the ETO.

Introduction

The Electronic Transactions Ordinance (ETO) (Cap 553) was enacted in January 2000 with a view to providing the legal infrastructure necessary for electronic commerce in Hong Kong. Following the Model Law on Electronic Commerce[1] adopted by the United Nations Commission on International Trade Law (UNCITRAL) in 1996, the main effect of the ETO is to give the same legal recognition to electronic records and digital signatures as to their paper-based counterparts. Now, two years after the ETO came into operation,[2] the Hong Kong Special Administrative Region (HKSAR) Government, based on its own commitment to review the Ordinance 18 months after its enactment, is conducting a public consultation on the ETO. The Government's view on the Ordinance is expressed in its *Consultation Paper on the Review of the Electronic Transactions Ordinance*[3] issued by the Information Technology and Broadcasting Bureau (ITBB) in March 2002. In the consultation paper, the ITBB proposes

* The authors are teaching staff of the Department of Computer Science and members of the Center for Information Security and Cryptography at the University of Hong Kong.

1 *See* http://www.uncitral.org/english/texts/electcom/ml-ecomm.htm.

2 The ETO has been in full operation since Apr 2000.

that the ETO be amended to accept personal identification numbers (PINs) to satisfy the signature requirement thereunder. The ITBB's argument is as follows:

> "The use of PIN for authentication has been widely tested in various types of market applications. With proper management, it can be considered for acceptance as a form of electronic signatures for satisfying the signature requirement under law in specified cases where the level of security offered by it is commensurate with the risk of the service involved, eg where there is already established relationship between the parties involved so that the PIN could be securely issued, used and verified; and where a secure system like the Electronic Service Delivery Scheme which provides strong encryption services for data transmission is used for making the electronic transaction ... We, therefore, consider that there is a case for the ETO to be amended and a new schedule added so that the Secretary for Information Technology and Broadcasting (the Secretary) may, by subsidiary legislation, specify in the new schedule legal provisions under which the use of PIN will be accepted for satisfying the signature requirement."[4]

The ITBB's argument is fundamentally flawed. Its proposal to accept PINs as a form of signature reflects a misconception about the basic requirements for a signature in general, and a lack of understanding of digital signature technology in particular. If the proposal were implemented, it would considerably damage the Government's endeavour to build a public key infrastructure (PKI) in Hong Kong and its efforts to promote electronic commerce based on such an infrastructure.

This article explains why, contrary to the ITBB's view, a PIN cannot serve as a signature for the purposes of the ETO. To appreciate this, one must start with the basic requirements for a signature and the principal features of the relevant technologies. As the analysis will demonstrate, once these requirements and features are properly considered and understood, it is clear why a PIN cannot serve as a signature under the ETO, whereas a digital signature based on PKI can.

Basic Requirements for a Signature

Despite the importance of a signature for commercial documents, the basic requirements for a signature are not expressly set out in legislation. However, if the relevant ordinances and case law are examined, three observations about the basic requirements can be made.

[3] *See* http://www.info.gov.hk/itbb/english/paper/index_n.htm.
[4] *Ibid.*, para 8.

First, the term "signature" is defined in the ETO to "include any symbol executed or adopted, or any methodology or procedure employed or adopted, by a person with the intention of authenticating or approving a record".[5] From this definition one immediately notes a basic requirement for a signature: namely, that it must indicate the signatory's approval of the document signed.

Second, although "signature" is not defined in any other ordinance, its meaning has been considered by the English courts in the context of various English statutes. In *Goodman v Eban*,[6] decided by the English Court of Appeal, Romer LJ approved a definition in *Stroud's Judicial Dictionary*[7] that "speaking generally a signature is the writing, or otherwise affixing, [of] a person's name, or a mark to represent his name, by himself or by his authority with the intention of authenticating a document as being that of, or as binding on, the person whose name or mark is so written or affixed".[8] This statement lays down another basic requirement for a signature, that it must represent the signatory so that the signed document carries the authority of the signatory.

Third, implicit in the statutory definition and case law mentioned above is a further requirement that a signature must not be forged or obtained by fraud, or else it is null and void. For bills of exchange, this requirement is implied in the Bills of Exchange Ordinance (Cap 19), which provides that "where a signature on a bill is forged or placed thereon without the authority of the person whose signature it purports to be, the forged or unauthorised signature is wholly inoperative".[9]

From these observations one can derive three basic requirements for a signature, whether the signature is handwritten or in digital form:

1 The signature must identify the signatory so that the document has the requisite authority of the signatory. This can be termed the "authorisation requirement".
2 The signature must indicate that the signatory has approved the contents contained in the document in their entirety. This can be termed the "approval requirement".
3 The signature must not be the product of any fraud, ie the signature must indeed be that of the signatory and must be applied by the signatory, or with his authority, to the document. This can be termed the "no fraud requirement".

These requirements will collectively be referred to as "the three requirements". They are consistent with the signature requirement suggested

5 Section 2(1).
6 [1954] 1 QB 550 (CA).
7 3rd edn.
8 Note 6 above, p 563.
9 Section 24.

by UNCITRAL in the Model Law on Electronic Commerce, from which the ETO and laws on electronic commerce in many other jurisdictions are derived.[10] The signature requirement is set out in the Model Law on Electronic Commerce in the following terms:

> "Where the law requires a signature of a person, that requirement is met in relation to a data message if:
> (a) a method is used to identify that person and to indicate that person's approval of the information contained in the data message; and
> (b) that method is reliable as was appropriate for the purpose for which the data message was generated or communicated, in the light of all the circumstances, including any relevant agreement."[11]

Although the requirement is phrased in the context of a "data message", it is clear that condition (a) therein corresponds to the authorisation requirement and the approval requirement, and that condition (b) contains the concept of the no fraud requirement. Indeed, the three requirements match even more closely the signature requirement under a more recent model law adopted by UNCITRAL, the Model Law on Electronic Signatures of December 2001.[12] Pursuant to this model law, an electronic signature meets the legal requirement for a signature if it is "as reliable as was appropriate for the purpose for which the data message was generated or communicated, in the light of all the circumstances, including any relevant agreement".[13] An electronic signature is considered to be reliable in this context if the following requirements are met:

> "(a) the signature creation data are, within the context in which they are used, linked to the signatory and to no other person;
> (b) the signature creation data were, at the time of signing, under the control of the signatory and of no other person;
> (c) any alteration to the electronic signature, made after the time of signing, is detectable; and
> (d) where a purpose of the legal requirement for a signature is to provide assurance as to the integrity of the information to which it relates, any alteration made to that information after the time of signing is detectable."[14]

[10] *See* for example the United Kingdom Electronic Communications Act 2000; the Australian Electronic Transactions Act 1999; and the Singaporean Electronic Transactions Act (Cap 88).
[11] Art 7(1).
[12] *See* www.uncitral.org/english/texts/electcom/ml-elecsig-e.pdf.
[13] Art 6(1).
[14] Art 6(3).

The similarity between these requirements and the three requirements is obvious: namely, requirement (a) corresponds to the authorisation requirement, requirement (d) corresponds to the approval requirement, and requirements (b) and (c) to the no fraud requirement.

Similarly, the three requirements are also consistent with the requirements for an "advanced electronic signature" under the European Union Directive on electronic signatures.[15] Under the Directive, such a signature is accorded the same legal status as a handwritten signature,[16] but it has to meet the following requirements as embedded in its definition:

> "'advanced electronic signature' means an electronic signature which meets the following requirements:
> (a) it is uniquely linked to the signatory;
> (b) it is capable of identifying the signatory;
> (c) it is created using means that the signatory can maintain under his sole control; and
> (d) it is linked to the data to which it relates in such a manner that any subsequent change of the data is detectable."[17]

Again, it is easy to see that requirements (a) and (b) in the definition correspond to the authorisation requirement, that requirement (d) corresponds to the approval requirement, and that requirement (c) is related to the no fraud requirement.

While the three requirements are closely related (particularly the authorisation requirement and the no fraud requirement) they remain separate. A simple example suffices to illustrate this point. A document satisfies the authorisation requirement by carrying a person's usual signature. Yet this does not imply that it also satisfies the no fraud requirement, as the signature may have been forged, or may have been obtained from the person on a separate occasion and attached to the document without consent. Nor does it imply that the document satisfies the approval requirement, as it may have been altered after the person has signed it.

In situations where a document is signed in the presence of all interested parties, the three requirements are normally satisfied without the need for further considerations. This is because each interested party is able to ascertain the true identity of the signatory and ensure that the document is properly signed. Moreover, it is usual practice in such circumstances for each interested party to be given a copy of the signed document, which serves as a

[15] Directive 1999/93/EC of the European Parliament and of the Council of 13 Dec 1999 on a Community Framework for Electronic Signatures: [2000] Official Journal L13/12.
[16] Art 5(1).
[17] Art 2(2).

safeguard against any alteration of the document by any party in the event of a dispute.

It is when a document is signed (or purported to be signed) in the *absence* of other interested parties that problems may arise. As the interested parties do not witness the signing of the document, they can only rely on the signed document itself to ascertain whether the document was indeed signed by the signatory free from fraud, and whether the signatory has approved of all the contents of the document. In short, the interested parties can only ascertain from the signed document alone whether all of the three requirements have been met.

Signatures on Paper Documents

Whether or not one is aware of it, many well-established methods for checking the "authenticity" of signatures on paper documents are in fact designed to ascertain if the signatures have met the three requirements. To illustrate this, some of the methods most often used in the conventional paper-based world are set out below.

Ascertaining Authorisation

To determine whether a signed paper document has met the authorisation requirement, the recipient of the document can check if there is external information which confirms that the signature represents the signatory. Most commonly, such external information is a record of the signatory's usual signature. If such a record is available, the recipient can compare the signature on the document with the one on the record. Where the two are substantially the same, the recipient can take the document as having the signatory's authority. This is the standard practice of banks in accepting and clearing signed cheques.

If the recipient has no record of the signatory's usual signature, as in the case of a first-time dealing with the signatory, the recipient can ask the signatory to sign the document in the presence of a trusted third party, such as a solicitor or a public notary. The third party is requested to verify the signatory's identity and also sign on the document to confirm that it has been properly signed by the signatory. Relying on such confirmation, the recipient can accept the signed document as having the requisite authority of the signatory.

Ascertaining Approval

To determine whether a signed paper document has met the approval requirement, the recipient can check if amendments have been made to the document, and, if so, verify that each amendment has the signatory's signature adjacent to it. Such signatures are explicit indications of the signatory's

approval of the amendments. If there is a signature missing, the recipient can reject the document on the ground that it has not been approved by the signatory in its entirety. Similarly, if a document is particularly important and consists of more than one page, the recipient can ask the signatory to sign on every page of the document and can reject the document if there is any page unsigned.

It is important to note the significance of having the signatory's signature adjacent to each amendment (or on every page) of the document. *Such signatures serve to "freeze" the contents of the document at the time the signature was made and render any amendments (or pages) that are subsequently included without authority (ie all those without the signatory's signature) immediately apparent to the recipient.* This ability on the part of the signatory to freeze a document is required to comply with the approval requirement. Without this ability on the part of the signatory, the recipient may never be able to tell whether the document she receives has been tampered with after it was signed.

Ascertaining No Fraud

To determine whether a signed paper document has met the no fraud requirement, the recipient can perform a variety of tests. These include: checking the pen strokes and special features of the signature to see if they are consistent with those of the signatory's usual signature; inspecting the paper to see if any deletion of or alteration to the signature has been made; and examining the position of the signature to see if it may have been taken from another document and attached to the document in question (as where the signature merely appears on an otherwise blank page at the end of a document). These methods have often been used in probate disputes concerning the authenticity of wills.

Generally, the most effective safeguard against fraud lies in having a sophisticated signature – the more sophisticated it is, the more difficult it is to forge. Ideally, the signature should be such that even those who have a record of it are unlikely to be able to forge it.

Alternatively, the recipient can adopt a similar arrangement as that for the authorisation requirement discussed above, by asking the signatory to sign the document in the presence of a trusted third party. The third party is requested to verify the identity of the signatory, to ensure that the document is properly signed, and to sign on the document as a witness thereto. If the document is executed in accordance with this arrangement, the recipient can accept it as being free from fraud.

Conditions for Compliance

From the above discussion, one can see the conditions for complying with the three requirements. Although the discussion focuses on signatures on paper

documents, the conditions apply equally to signatures on any kind of documents. It is thus worth summarising them:

1 to comply with the authorisation requirement, there must be information available to the recipient of the signed document to confirm that the signature represents the signatory;
2 to comply with the approval requirement, the signatory must have the ability to "freeze" the contents of the document at the time it is signed; and
3 to comply with the no fraud requirement, the signature must be made as sophisticated as is practicable so that even those who have a record of the signature are unlikely to be able to forge it.

Electronic Documents

Unlike paper documents, electronic documents are made up of sequences of bits[18] and reside on digital media. As such, there are two characteristics of electronic documents that make them fundamentally different from paper documents.

First, electronic documents can be copied without any loss of quality. Hence there is no concept of "original copy" for electronic documents, as any copy made is identical to, and is just as good as, the original.

Second, electronic documents are seamless. They can be altered easily and without leaving a trace. If done carefully, any part of an electronic document can be cut out and inserted into another electronic document without the change being noticeable. The resulting documents will appear just as seamless as any other electronic document.

These two characteristics render the forging of electronic documents much easier than that of paper documents. As there are generally no discernible differences between a forged electronic document and a genuine electronic document, anyone unfamiliar with the digital technology could be deceived. Furthermore, because there is no concept of "original copy" for electronic documents, methods for ascertaining the validity of signatures that rely on the existence of original copies (such as those used for paper documents discussed earlier) cannot be applied to electronic documents.

To provide the necessary technological infrastructure for signing electronic documents, one needs a mechanism for generating signatures on electronic documents that comply with the three requirements, and a method that

[18] "Bit" is short for *b*inary digi*t*, the smallest unit of information in the digital world. A bit can hold only one of two values: 0 or 1. More meaningful information is represented in larger units obtained by combining consecutive bits. For example, a byte is composed of eight consecutive bits, which is the unit for representing an English alphabet or an Arabic numeral.

enables anyone to determine the validity of those signatures. It is for these purposes that digital signature technology has been devised. Although the technology is now well known, it is instructive to review why it can fulfil these purposes, in order to see why PIN technology cannot.

Digital Signatures

Central to digital signature technology is the notion of the "digital signature", which is defined in the ETO as:

> "an electronic signature of the signer generated by the transformation of the electronic record using an asymmetric cryptosystem and a hash function such that a person having the initial untransformed electronic record and the signer's public key can determine –
> (a) whether the transformation was generated using the private key that corresponds to the signer's public key; and
> (b) whether the initial electronic record has been altered since the transformation was generated."[19]

This rather technical definition reflects the process of generating digital signatures for electronic documents. Briefly, the process consists of two steps:

1 A standard mathematical function known as a "hash function" is applied to the contents of the electronic document to produce a "hash value". This hash value (a sequence of 160 bits if it is produced by the standard SHA-1 hash function)[20] represents a digest of the contents of the document.
2 The hash value is encrypted (ie scrambled) by the signatory using his unique private key. The value can only be decrypted (ie unscrambled) by the corresponding public key.[21] *This encrypted hash value is the "digital signature" for the electronic document.*

[19] Section 2.

[20] "SHA" stands for "Secure Hash Algorithm". Another standard hash function is MD5 ("MD" is short for "Message Digest"), which produces hash values of 128 bits. For more information about these standard hash functions, *see* Menezes, van Oorschot and Vanstone, *Handbook of Applied Cryptography* (New York: CRC Press, 1997), Ch 9.

[21] Private and public key pairs are generated based on the RSA Public-Key Cryptosystem invented by three professors at the Massachusetts Institute of Technology in 1977. The name RSA was derived from the names of the three inventors, Rivest, Shamir and Adleman. The RSA Public-Key Cryptosystem is an "asymmetric cryptosystem" as it involves the use of key pairs and it is computationally infeasible to deduce one key of a key pair from the other, ie a private key cannot be deduced from the corresponding public key, and vice versa. For more details on the RSA Public-Key Cryptosystem, *see* Rivest, Shamir and Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems" (1978) 21 *Communications of the Association for Computing Machinery* 120. *See* also US Patent 4,405,829.

From both the legal and technological points of view, the two steps above are significant. Because the hash value for a document is in effect a fingerprint for the document,[22] and any alteration to a document will result in a different hash value, the effect of step (1) is to freeze the contents of the document at the time of creating the digital signature. By encrypting the hash value using the signatory's private key (a unique key known only to the signatory and which is virtually impossible for anyone to derive or guess) the effect of step (2) is to link uniquely the digital signature to the signatory, ie the owner of the private key. As will be seen, these two steps are crucial for generating signatures that comply with the three requirements.

The above process for generating digital signatures also creates an important property in the digital signature which distinguishes it from handwritten signatures. This property is often overlooked. While a handwritten signature is an attribute of the signatory alone, the digital signature is an attribute of a combination of the signatory's private key and the contents of the electronic document. In other words, a handwritten signature is only "signatory-specific", whereas the digital signature is both "signatory-specific" and "document-specific". *It follows that while one can expect handwritten signatures of the same signatory on different documents to look substantially the same, one cannot expect this to be the case for digital signatures.* Indeed, the digital signatures for two different documents are always different, even when they are produced by the same signatory using the same private key. With the slightest change in the contents of an electronic document (for instance, changing one occurrence of "e" to "i"), an entirely different digital signature will ensue.

Compliance with the Three Requirements

The compliance of the digital signature with the three requirements and the verification of such compliance is built into the process for sending and receiving digitally signed documents. In this process, the signatory sends a digitally signed document to the recipient, who decides whether or not to accept it based on the digital signature. This is done as follows.

After generating the digital signature for an electronic document, as described earlier, the signatory sends the electronic document to the recipient together with the following information:

[22] The hash value for each document is almost certain to be unique. For example, if one uses the standard SHA-1 hash function, the probability of two different documents resulting in the same hash value is one in 2^{160}, ie less than one in one trillion trillion trillion trillion, or virtually nil.

1 the digital signature; and
2 a digital certificate issued by a recognised certification authority (ie a trusted third party)[23] which confirms the signatory's identity and which contains the signatory's public key.

Upon receiving the digitally signed document, the recipient performs the following steps to decide whether or not to accept the document.

First, the digital signature is decrypted using the signatory's public key contained in the digital certificate. This recovers the original hash value encrypted by the signatory.

Second, the standard hash function used by the signatory, known to the recipient, is applied to the contents of the document to recompute the hash value. This value is compared with the original hash value recovered in the first step. If the two hash values are the same, the document is accepted; otherwise it is rejected.

While not immediately obvious, the steps performed by the recipient actually serve to ascertain whether the digital signature has met the three requirements. *What is crucial here is the criterion for accepting a digitally signed document: namely, that the hash value recomputed by the recipient must be the same as the original hash value recovered from the digital signature.* If this criterion is satisfied, it means that:

1 The document has not been altered since it was digitally signed. Otherwise the recomputed hash value (based on the altered document) would be different from the original hash value recovered from the digital signature (based on the document at the time it was digitally signed).
2 The digital signature was created by the unique private key corresponding to the public key contained in the digital certificate attached to the document. If this were not the case, the recipient would not have been able to recover the original hash value from the digital signature and to find that it matched the recomputed hash value.

Taken together, these two facts imply that the three requirements are met. It is not difficult to see why.

First, fact (1) precisely means that the digital signature satisfies the approval requirement.

Second, fact (2) means that the digital signature was created by, or with the authority of, the owner of the private key, whose identity is confirmed by

[23] In Hong Kong, the Postmaster General is a recognised certification authority by virtue of s 34 of the ETO. At the time of writing, three more certification authorities have been recognised by the Director of Information Technology Services: Digi-Sign Certification Services, Joint Electronic Teller Services and HiTrust.com.

the digital certificate issued by a trusted third party. This satisfies the authorisation requirement.

Third, because of the asymmetric nature of the public key cryptosystem,[24] it is infeasible to derive the corresponding private key from a public key computationally, let alone by a "blind guess".[25] It follows that it is virtually impossible that the private key could have been forged by anyone, including those who know about the public key. Hence fact (2) also means that the no fraud requirement is satisfied.

It is important to note that although electronic documents are seamless, one cannot forge a digital signature by a "cut-and-paste" operation (ie taking the signatory's digital signature from one document and pasting it onto another). Such a signature would be rejected outright, as the original hash value recovered therefrom would be based on another document and would therefore not match the hash value recomputed from the document in question.

A close study of digital signature technology reveals the fundamental reason why the digital signature can comply with the three requirements. It is because the technology has built into it the following elements for compliance:[26]

1 with regard to the authorisation requirement, compliance is achieved by the digital certificate accompanying the document, which provides the information for confirming the signatory's identity;
2 with regard to the approval requirement, compliance is achieved by computing the hash value for the electronic document, which serves to freeze the contents of the document at the time of creating the digital signature; and
3 with regard to the no fraud requirement, compliance is achieved by the inherent improbability of anyone deriving the signatory's private key from the corresponding public key.

All of these elements are indispensable to the eligibility of a digital signature as a signature. If there is any element missing, a digital signature will not be able to satisfy all of the three requirements.

[24] *See* n 21 above.

[25] In theory, one should be able to find the private key by using a "brute force" approach of trying each and every possible value that may lead to deducing the private key. But the time required for this would be astronomical. Using the current standard 1024-bit RSA Public-Key Cryptosystem as an example, the brute force approach would have to perform a mathematical process known as "factorisation" by trying all possible values within the range 0 to 2^{512}, a total of roughly 10^{154} (ie 1 followed by 154 zeroes) possible values. Even if one had a machine which could try one billion different values per second — more powerful than today's top-end desktop computers — it would take more than 10^{137} (ie 1 followed by 137 zeroes) years to exhaust all possible values in order to deduce the private key. This renders the approach practically infeasible. More sophisticated methods would employ highly advanced mathematical techniques to reduce the time for deriving the private key. But despite efforts in this regard, at the time of writing there is no reported case where the 1024-bit RSA Public-Key Cryptosystem has been successfully compromised anywhere in the world.

[26] *See* "Conditions for Compliance" under "Signatures on Paper Documents" above.

In practice, digital signature technology has two disadvantages that need to be overcome. First, because the technology requires a digital certificate which has to be obtained from a recognised certification authority, it is not entirely convenient to the end user.[27] But overcoming this is essentially a matter of making digital certificates cheaper and easier to obtain. As the market sees more transactions conducted using digital signatures, there should be more incentive for the public to adopt the technology. Second, as private keys are almost impossible to memorise, they are invariably stored on computers or other digital devices. As such, they are prone to security attacks. If a private key falls into the wrong hands, the unique link between the key and the key owner is immediately compromised. But this kind of risk is not unique to the digital signature; it is in fact an inevitable and perennial risk that all security technologies must face. For the digital signature, as with the password and PIN, the solution lies in adopting sound key management measures and polices, at both the personal and corporate levels, which ensure the security of private keys.

Personal Identification Number

Unlike the digital signature, the PIN does not involve the use of a key pair. The concept of the PIN is akin to that of an identity card. The idea is that each individual is given a unique PIN, and any person who is able to tender a valid PIN is deemed to be the owner of that PIN (or at least have the authority of the owner of the PIN). Thus, right from the beginning the PIN is designed for authorisation. The main advantage of the PIN is ease of use, as it does not require obtaining a digital certificate from a third party and using a private key that is difficult to remember. Because of this, the PIN has long been used for the purpose of authorisation in a variety of applications, most notably banking transactions and computer logins.

Depending on the method used, a PIN may be checked by its recipient, either in its original form or in hashed form (ie checking is performed on a hash value generated for the PIN instead of the PIN itself, as is done in the UNIX operating system in respect of passwords). In any event, the recipient must maintain a database containing all the authorised PINs in either their original or hashed form, together with information about their corresponding owners. Upon receiving a PIN, the recipient will check to see if the PIN or its

[27] This is one of the reasons why digital signatures have not been widely used in Hong Kong. In an article published in May 2002, it was reported that Hongkong Post had issued about 50,000 personal and corporate digital certificates, which was only one-tenth of what it expected when it launched its service in early 2000: "New Certification Authorities Recognized", *Computerworld Hong Kong*, 10 May 2002, p 4.

hash value is stored in the recipient's database. If it is, then the PIN is deemed to be entered by the PIN owner.

Once a PIN owner has logged onto the recipient's system this way, she does not need the PIN again to perform any action during the login session. There is no requirement that the PIN must be contained in or attached to any electronic document or message submitted by the PIN owner during the login session. The PIN owner will only need to supply the PIN the next time she starts a new login session. The recipient will, for accounting and security reasons, keep a record of when a PIN owner logs in and when she logs out (commonly known as a "session record"). In case of dispute, such records will serve as evidence as to whether or not an act was performed by a particular PIN owner at a particular time.

The way PINs work clearly shows that they are designed solely for the purpose of authorisation and were never intended to play the role of signatures. Indeed, PINs can only satisfy the authorisation requirement, not the other requirements for a signature. A moment's thought reveals the reason why. Since electronic documents submitted in a login session do not require a PIN, there is nothing to prevent a dishonest recipient of a PIN from fabricating an electronic document and creating a session record alleging that the document was submitted by a PIN owner during that login session. To any outsider, including the PIN owner, there is no easy way to discern such a fabricated document and session record from genuine ones.

Furthermore, even if a PIN mechanism requires PIN owners to attach PINs or their hash values to electronic documents submitted during a login session, a dishonest recipient is still capable of practising the same fraud. This is because the attachment of a PIN to a document does not "freeze" the contents of the document. As the recipient has a database containing all the PINs or their hash values, he can extract any PIN or its hash value from the database and attach it to a fabricated electronic document at any time if he so wishes. Such fabricated documents are indistinguishable from genuine documents.

The ability of the recipient to forge electronic documents clearly violates the no fraud requirement. In stark contrast, it is virtually impossible for the recipient of a digitally signed document, and indeed anyone other than the owner of the private key, to fabricate an electronic document alleging that it was produced by the owner of the private key.

A closer examination of PIN technology reveals its inherent primitiveness and lock of security as compared to the digital signature:

1 the PIN does not involve the concept of an asymmetric key pair, whereas the digital signature is built upon such a concept; and
2 the PIN is not attached to any document submitted by the PIN owner, and even when it is it does not "freeze" the contents of any document

it is attached to, whereas the digital signature always "freezes" the document on which it appears.

It follows that the PIN cannot satisfy the approval requirement either. Since an electronic document submitted in a login session is not required to contain a PIN, the document can be altered without the change being detected. In addition, even where a PIN is contained in such a document, the PIN does not freeze the contents of the document. Hence the document can also be altered without being detected.

Because the PIN does not satisfy all of the three requirements, it can never be used as a fully-fledged signature. It can indicate authorisation, but not approval of contents or absence of fraud. The fundamental characteristics of PINs are not changed, however extensively they may have been used in market applications. Unfortunately, this is precisely what the ITBB has failed to appreciate in the Government's consultation paper. As a result, its argument in proposing that the PIN be accepted as a form of electronic signature is fundamentally flawed.

The truth of the matter is that, owing to the inherent deficiencies of PINs, any arrangement purporting to treat PINs as "signatures" will always place PIN owners in a vulnerable position. A PIN owner can only trust that a recipient, with the full capability to forge electronic documents and fabricate session records, will act honestly and not betray her trust. This is the kind of trust that bank customers who use PINs for banking transactions have always placed in the banks. This is also the kind of trust that taxpayers have to place in the Government if passwords are to be accepted as signatures for electronic tax returns, as proposed in the Inland Revenue (Amendment) (No 2) Bill 2001,[28] which is mentioned in the consultation paper.

Such trust in the recipient is fundamental to the use of PINs. It cannot simply be imposed or assumed by legislation, but can only arise as a matter of contract between the PIN owner and the recipient (as in the case between a bank customer and the bank). In this regard, the ITBB's suggestion to use "a secure system like the Electronic Service Delivery Scheme which provides strong encryption services for data transmission"[29] is wholly irrelevant. Such secure communication channels can only protect the PIN owner against eavesdroppers, but not against the recipient. However secure a communication channel is, it can never prevent a dishonest recipient from forging or altering electronic documents and alleging that such documents were submitted by the PIN owner.

[28] Section 2(b) of the Bill proposes to add a new s 2(5) to the Inland Revenue Ordinance (Cap 112) which reads: "In this Ordinance, a reference to the act of signing a return required to be furnished under this Ordinance includes a reference to the adopting of — (a) a digital signature (supported by a recognized certificate and generated within a period during which the certificate is valid); (b) a password; or (c) any other signing device, for the purpose of authenticating or approving the return."

[29] *See* n 4 above.

It is in this respect that the digital signature shows its greatest strength, namely, that its use does not depend on any trust in the recipient. On the contrary, digital signature technology is designed precisely on the assumption that the recipient is not to be trusted; the technology guarantees that it is virtually impossible for the recipient of a digitally signed document to forge a digital signature of an owner of a private key.

Conclusion

Whether a signature is handwritten or digital in form, it must satisfy three basic requirements in order to give legal effect to documents that carry the signature. These requirements can be termed the "authorisation requirement", the "approval requirement" and the "no fraud requirement". They are consistent with the signature requirements under the UNCITRAL Model Law on Electronic Commerce, the UNCITRAL Model Law on Electronic Signatures and the European Union Directive on electronic signatures.

In the paper-based world, there are well-established methods to ascertain whether a signature has met these requirements. In the digital world, digital signature technology has also built into it all the necessary elements for compliance with the requirements. In contrast, PIN technology is designed solely for the purpose of authorisation. As shown, despite its relatively long history and ease of use as compared to the digital signature, PINs can only satisfy the authorisation requirement, not the approval requirement or the no fraud requirement. Because of these inherent deficiencies, PINs do not qualify as fully-fledged signatures. Hence, the ITBB's proposal that the ETO be amended to accept PINs as satisfying the signature requirement thereunder is fundamentally flawed and should be firmly rejected.

Although there is no formal proposal yet, there are suggestions that "biometrics" – the emerging field of technology devoted to the identification of individuals based on their biological traits, for example retinal or iris scanning, fingerprint recognition and face recognition – can be used for signing electronic documents. Unlike PIN technology, which relies on randomly generated sequences of numbers and / or letters to represent individuals, biometrics is based on each individual's unique biological traits that cannot be forged. Thus, biometrics is more closely connected to the individual than PINs are, making it well-suited for direct identification of individuals rather than just for ascertaining authorisation. However, like PINs, the use of biological traits as a form of electronic signature also suffers from two loopholes:

1 the attachment of biological traits to a document does not freeze the contents of the document; and

2 the recipient of documents thus signed must, for the purpose of verifying the signatures, maintain a database containing the biological traits of individuals in either their original form or some other form.

As explained in the context of the PIN, these two loopholes will also render biometrics incapable of satisfying the approval requirement and the no fraud requirement. Accordingly, unless these loopholes are closed, biometrics alone cannot be used to sign electronic documents.

As technology stands, the digital signature based on PKI is the most secure and practical solution for signing electronic documents. Experience in other countries also shows that the digital signature is the technological foundation for electronic commerce. The ETO, in giving legal recognition to the digital signature, has provided the necessary legal infrastructure for electronic commerce in Hong Kong. It may take some time for digital signature technology to mature and become widely used in the community. Even so, the legal infrastructure for electronic commerce already established by the ETO must not be undermined by extending legal recognition to forms of electronic signatures that are less secure than digital signatures, let alone those that do not even meet the basic requirements for a signature.