Encryption – use and control in E-commerce

by Robert Bond

Robert Bond

The author describes how cryptography can be used to address modern business requirements.

WHAT IS CRYPTOGRAPHY?

Cryptography is the art of secret writing and has been used to conceal the contents of messages from potential adversaries for thousands of years. In ancient Greece, the Spartan generals used a form of cryptography so that they could exchange secret messages. The messages were written on narrow ribbons of parchment that were wound spirally around a cylindrical staff called a *scytale*. After the ribbon was unwound, only a person who had a matching cylinder of exactly the same size could read the writing on it.

Nowadays thankfully, the cryptographic methods used in secure e-commerce are considerably more sophisticated, and are used to support more than just the confidentiality of a message. Modern cryptography provides a basis for addressing many business requirements, including integrity protection, authentication and hence accountability, protection against repudiation, and detection of unauthorised copying.

Cryptography can in some ways be compared to the lock used on the door of a house or car and makes use of two components to function properly:

- the algorithm, which for the purpose of this discussion can be considered as being akin to the lock itself; and,
- the key, which is used to operate the lock.

Some normal, everyday locks are more easily broken or picked than others. Some locks have a more secure design than others, but if the key is left in an obvious place (under the doormat?), how effective is the lock? If the lock is constructed of high strength material but has a relatively simple design, then it can easily be picked. Conversely, if the design of the lock is good, but it is constructed poorly or out of weak materials, then no matter how sophisticated or strong the key, the lock can easily be broken.

To ensure a strong and effective lock, the design, the material from which it is constructed and the key must satisfy criteria appropriate to the application to which it will be put. The key must also be protected from unauthorised use. A weakness in respect of any of these criteria, or a lapse of security in respect of the key, will render the whole set-up useless. It is the same with cryptography: the algorithm must be of a good strong design, the implementation of the design must be done well and without flaws, and both must be capable of withstanding attacks even when the attacker knows the design and implementation in detail, *(remember, no security through obscurity)*. Just as with physical keys, cryptographic keys have to be protected from unauthorised use or the security of the whole set-up is compromised.

There are many excellent books that provide extended tutorials on cryptography, so only the salient points will be covered here. Modern cryptography falls into two main camps: symmetric or secret key cryptography, and asymmetric or public key cryptography. The main difference between the two types is that the former uses the same single key to both encrypt and to decrypt, while the latter uses one key to encrypt and another to decrypt.

When using secret key cryptography to protect a message or some other type of exchange, both the originator and the recipient of the message need to have access to the same key that is used for both the encryption and decryption operations. So somehow, that key needs to be distributed to where it is needed. Here we have a difficulty and in the past this was solved by distributing keys by what are called 'out of band' methods, such as, for instance, delivery by hand. This works well for low volumes of encrypted information, where the key may not need to be changed very often or for cases where the number of individuals with whom we wish to communicate is small (and therefore a small number of keys). When we wish to begin supporting secure e-commerce over the internet for instance, where the number of potential participants is unlimited, the key distribution and management problems associated with secret key cryptography really begin to surface. This is especially true in modern internet-based e-commerce situations, which can be dominated by communications with individuals with whom we have had no prior relationship.

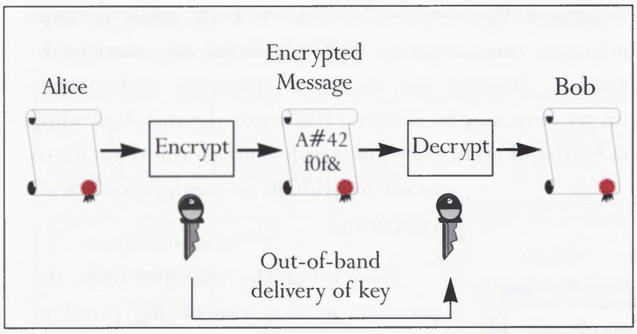


Figure 1 Symmetric or Secret Key encryption

Thankfully, we have a solution to this problem of securely distributing keys in public key cryptography. In public key cryptography, two keys are generated *for each individual*: a private key, (not to be confused with the secret key described earlier), and a public key. The individual for whom the key pair is generated must protect the private key from others (i.e. must keep it private). That person is free however, to distribute the corresponding public key as freely as he or she wishes, and via any secure or completely insecure mechanism.

So in the case that Bob wishes to send a confidential message to Alice, Bob retrieves Alice's public key (perhaps it is on a web server, or on other public access channels that we will discuss later). Bob can then encrypt the message using Alice's public key, and send the resulting encrypted message to Alice. Only the person with the corresponding private key is able to successfully decrypt the message, (in this case Alice), and so only Alice can decrypt and read the contents of the original message. Similarly, Alice could send a confidential message to Bob, (by using Bob's public key to encrypt the message so that only Bob can later decrypt the message using his private key).

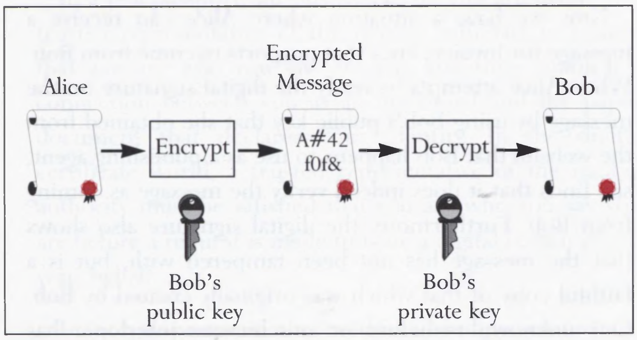


Figure 2 Public Key encryption

From the above illustrations, by just switching to public key cryptography, it would seem that we have solved the problems of key management and distribution that are associated with secret key cryptography. Apart from hugely reducing the number of keys that are required to support the secure exchanges of any given population of participants, we have also provided a workable method of distributing keys using insecure channels, but without compromising the security of our communications.

It would seem that since the advent of public key cryptography, we would have no further use for secret key cryptography. Unfortunately this is not the case, since public key cryptography is much slower to carry out than its secret key counterpart. If we were to rely solely on public key cryptography for protecting our personal or business exchanges, they would be reduced to a crawling pace. Consequently, it is normal practice to continue to use symmetric or secret key cryptography for encrypting bulk data such as the contents of an email, document, contract or invoice, etc. Public key cryptography is then used to securely deliver the symmetric or secret key to the recipient, where it is needed to decrypt the bulk data item. Under this scheme, randomly generated symmetric keys can be used for each session. In many electronic payment schemes and web access schemes for instance, these symmetric keys are also called *session keys*, for that very reason.

WHAT ELSE CAN BE DONE WITH CRYPTOGRAPHY?

So far we have seen how private and business exchanges can be made confidential using cryptography. The very same technologies however, can be used to provide integrity protection and proof of origin. Remember that in public key cryptography, two keys are generated for each individual: a private key and a public key, and that for confidentiality protection, the recipient's private key is normally used to decrypt a message that has previously been encrypted with the corresponding public key.

Now let us suppose that again, Bob wanted to send a message to Alice, but this time he is not interested in keeping the message confidential, but is certainly interested in enabling Alice to determine if the message definitely came from Bob, and not from someone masquerading as Bob. In this case, Bob could encrypt the message with his private key, and then send the resultant encrypted message to Alice. Given the properties of public key cryptography, we know that if a message is encrypted with the public key, only the corresponding private key can be used to decrypt it. The converse is also true; if a message is encrypted with the private key, only the corresponding public key can be used to decrypt it. Therefore in this situation, Alice now knows that if she can successfully decrypt the message purporting to come from Bob, using Bob's public key, then the message could only have been encrypted using Bob's private key, and therefore the message must have come from Bob. In real life, because of the poor performance of public key cryptography, these operations are normally performed on a piece of data that is much smaller than the bulk data item we are sending. This smaller piece of data, which is essentially a very large number, is called a *hash or message digest*, and has properties such that it is:

- infeasible to determine the input message from its digest;
- infeasible to find an arbitrary message that will produce a particular specified digest; and
- infeasible to find two different messages that will produce the same digest.

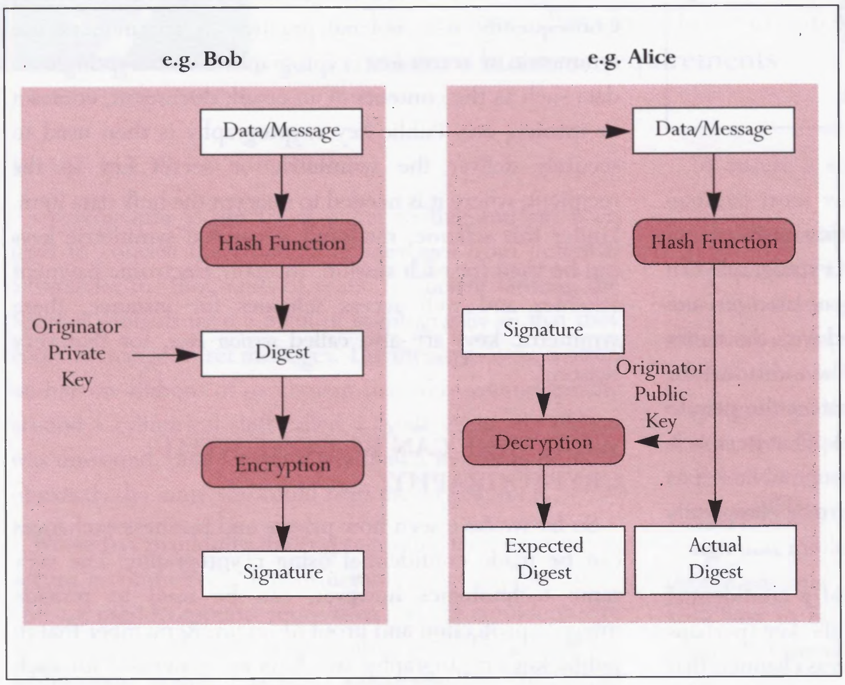


Figure 3 Digital Signature

What this means is that if the message were to be changed by even one character or one bit, the message digest would suffer drastic change, (perhaps as many as half of the bits in the digest might be changed). By using this message digest instead of the whole message, Bob merely has to encrypt the message digest using his private key, and then send it along with the message itself. Alice then re-computes the message digest from the message, decrypts the message digest that was sent to her by Bob (using Bob's public key), and compares the two values. If they are the same, Alice can be confident that the message did in fact come from Bob. We have in fact, applied what is known as a *digital signature* to the message (specifically in this case, Bob's digital signature).

Another property of this process is that if the two message digests are the same, Alice can also be confident that the message was not changed en-route. Hence we have also provided some protection of the integrity of the message. This is useful in cases where knowledge of whether the message has been tampered with is important, (for instance, altering the amounts on an invoice or payment).

MAKING THE TRANSITION FROM SECURITY TO TRUST

Although security is essential to be able to engage in electronic commerce, security by itself is not what we need to achieve. We need to be able to trust the infrastructure on which we depend to facilitate our private and business exchanges. We need to be able to trust other perhaps unknown, unseen parties with whom we may want to do business. Security and secure e-commerce technologies can go some way to enabling this trust; the rest depending on legal and regulatory safeguards, and on the reputations of the individuals or commercial brands concerned.

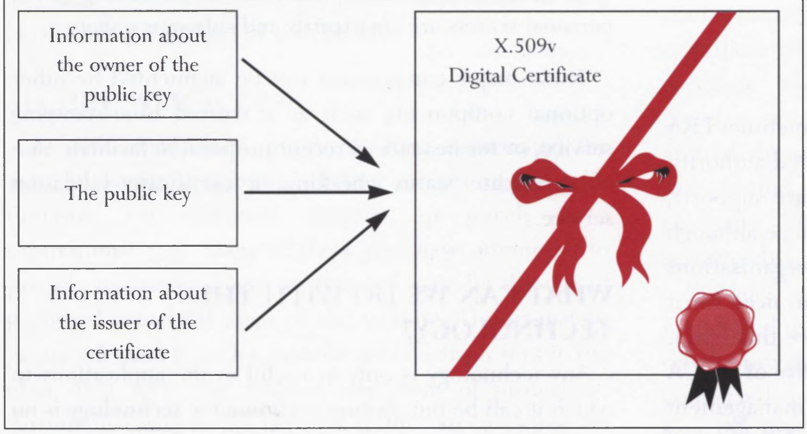
Remember the example from the previous section concerning proof of origin. In this example, Alice satisfied herself that the message originated with Bob because she was able to decrypt the encrypted message digest or hash, by using Bob's public key, (otherwise known as verifying Bob's digital signature). *But was it Bob's public key?* How can we be sure? As was mentioned earlier, public keys have the valuable property of being capable of distribution via any non-secure mechanism, such as merely being published on a website, or being emailed around to various interested parties. But let's imagine that some malicious or mischievous person wanted to masquerade as Bob. All he or she would have to do would be to generate a new key pair, and somehow distribute the public key of this new key pair with the announcement that it was in fact Bob's public key. Normal email has been found reasonably easy to fake, and successful attacks on websites to replace content are not uncommon, so it should be possible in many cases to substitute Bob's public key with another.

Now we have a situation where Alice can receive a message (or invoice, etc.) that purports to come from Bob. When Alice attempts to verify the digital signature on the message by using Bob's public key that she obtained from the website that Bob happens to use as a publishing agent, she finds that it does indeed verify the message as coming from Bob. Furthermore, the digital signature also shows that the message has not been tampered with, but is a faithful copy of that which was originally created by Bob. Our unknown malicious or mischievous interloper has successfully masqueraded as Bob.

Here is a clear situation where we have elements of very strong security, but very little trust. We can protect messages to make them tamper-resistant, or to make them secret; we can even provide irrefutable proof that the message was generated using a particular private key when digitally signing the message. *But we cannot be sure to whom the key belongs.*

To make this transition from security to trust, we need to introduce a new element into our burgeoning infrastructure: the X.509 digital certificate, or to be precise with respect to current developments: the X.509v3 (for version 3) digital certificate.

Figure 4 The X.509v3 Digital Certificate


Information about the owner of the public key
The public key
Information about the issuer of the certificate
X.509v
Digital Certificate

The sole purpose of this certificate is to bind an entity such as a person, company, department, machine, or software agent, etc., to a public key. It aims to do this indivisibly and in a manner that can be trusted. In order to do this, it borrows an idea that has been around for quite some time in the paper-based world: the trusted authority or trusted third party.

In simple terms, the X.509v3 digital certificate is the electronic commerce world's analogue of the passport. Like the passport, it is issued by a trusted authority and binds you as an individual to an identity that can be recognised and verified by other agencies (the public key). On issuance, it confers certain rights and obligations on you according to policies exercised by the issuing authority.

In a real passport, various checks on you are made by a trusted representative of the issuing authority to ensure that you are who you say you are, and thus establish a connection between you as an individual and the paper document that declares your identity. In the digital certificate world, a trusted representative of the issuing authority must be satisfied that you are who you say you are before a request is made to issue a digital certificate on your behalf.

In a real passport, the methods used to ensure the integrity of the binding between you and the paper identity are such things as watermarks, seals, special paper and ink, etc. In the digital certificate world, the method used to ensure the integrity of the binding between an individual or other entity and the public key is the digital signature of the issuing authority.

Because the X.509v3 digital certificate uses and supplies, relying parties with the tools of cryptographic technology, it provides you with the ability to digitally sign documents or transactions, or to verify the signatures of others. It enables you to make documents or transactions only readable by those that you designate.

MANAGING IT ALL

We've seen how we can create a digital certificate so that we can inextricably bind a public key to a recognisable and accountable identity such as a person, a company, a software agent, or a machine. We've also seen how these digital certificates mirror the world of paper certificates such as passports, in some ways. In this paper-based world, we are used to the existence of trusted agencies (such as the passport office or the credit card company), to look after the issuance, revocation and general management of these certificates. In the burgeoning world of electronic commerce, in order to manage the huge number of digital certificates that might be in circulation, similar trusted agencies must be created.

In this brief tour of the technical underpinnings of secure electronic commerce, we will not go into any great detail concerning agencies intended to manage digital certificates, but will briefly list them, and their functions.

PUBLIC KEY INFRASTRUCTURE

The first thing to know is that the infrastructure as a whole that is intended to issue, manage and facilitate the use of digital certificates, is called a Public Key Infrastructure or PKI for short. It is a term that you are likely to hear many times if you become involved in secure electronic commerce. A PKI consists of the following components.

The Certificate Authority

Otherwise abbreviated as CA, this is a trusted authority, embodied in software (with possible hardware support), that is responsible for certificate management operations on behalf of a community of certificate users, (or as the American Bar Association describes them – relying parties). These relying parties could be people, file-servers, web-servers and the like, business applications, mobile software agents, or whatever is required to be able to communicate with confidentiality, integrity, authentication, etc. The range of certificate management operations undertaken by the CA encompasses certificate issuance, renewal, revocation, suspension, retirement, and archival. Under some circumstances, the CA may also be responsible for actually generating the key pairs used in the processes outlined earlier. Its primary function however, is to act as a trusted authority that vouches for the binding between an identity and a public key, and hence vouch for the validity of the electronic identity (instantiated in the public key) of any of its relying parties. In short, the CA is the entity that can be trusted to say to anyone *'This is Bob's public key'*. Consequently, the keys that the CA uses for signing certificates (to ensure the binding between identity and public key) should be regarded as the *Crown Jewels* of the infrastructure and should be very strongly protected. If these were ever to fall into the wrong hands, certificates could be forged quite easily.

Optionally, the Registration Authority

This is otherwise abbreviated as RA, (or sometimes LRA – the L meaning Local). Again this is a trusted authority, embodied in software (with possible hardware support), but this time it is an optional component, since although it is commonly quite useful, some organisations (particularly small ones) may have little or no need for it (whilst other organisations may require more than one). Its main role is to act as a trusted representative of the CA to which the CA can delegate some management functions. These functions being the registration of individuals or other entities for inclusion into the community of certificate users, requests for revocation, suspension, or update. Commonly, the RA software is used by an authorised individual from the community being served, (such as someone from the personnel department of a company), whose job it is to ensure that sufficient proof of identity and eligibility is produced before a certificate is issued. Consequently, rather than the CA, which is a central resource, the RA is usually located where it will be most useful (usually near to the community of certificate holders that it serves). Hence the RA is also known as a Local Registration Authority.

The Directory

Like a telephone directory in which the telephone numbers of subscribers are published, the directory associated with a public key infrastructure is the place where subscribers' digital certificates are published. Remember that these digital certificates contain the subscribers' public keys, and so the directory is the place to look if you want to send confidential messages to a subscriber, or if you want to check his or her digital signature. The benefit of using directories for publishing certificates, as opposed to any number of alternative methods (such as flat files, various web page formats, etc.), is that directory services increasingly can be accessed by a standard mechanism that facilitates automatic access and processing in business software. This is called the Lightweight Directory Access Protocol or LDAP. Directories are also used to publish notifications of certificate revocations and suspensions, and so this is also the place to discover if any particular certificate is still valid.

The Personal Security Environment

This is a term used to describe a variety of methods employed to protect personal secrets. In particular, the secrets with which we are currently most concerned are the private keys that have been generated as one half of the key pairs used in public key infrastructures, and hence secure electronic commerce. Some more familiar terms that describe individual methods of protecting these personal secrets are smartcards and software wallets.

These major components can be augmented by other optional components such as a trusted time-stamping service, or for instance, a recent proposal to facilitate ease of certificate status checking: a certificate validation service.

WHAT CAN WE DO WITH THE TECHNOLOGY?

Any technology is only as useful as the applications to which it can be put. Secure e-commerce technology is no exception, so we must be able to demonstrate improved ways of doing things or enable entirely new and useful things to be done that could not have been accomplished before. In what follows, we will briefly examine some of the higher level facilities that can be built using secure e-commerce technology. These are not by themselves what might be called business applications, but combination and co-ordination of such facilities by business-specific application code and processes, can build powerful new business applications.

SECURE EMAIL

Many millions of business and individuals have come to rely on email as a cheap and efficient form of communication that, in the main, works well without regard to differences in location or time zone. It is relatively easy to use and the benefits are generally well understood. Consequently the proportion of individuals and businesses becoming email-enabled is growing at a tremendous rate. Only now are some of the email converts beginning to appreciate some of the risks associated with the use of email.

(1) Email can easily be forged to appear to come from someone else.

(2) As a consequence of the previous point, anyone can assert that they had never sent some particular email.

(3) Alternatively, people can assert that they never received some particular email.

(4) Email addresses are not sufficiently well bound to a real identity, so the email recipient may not be the intended recipient, but someone masquerading as such.

(5) Email can be modified in transit without anyone being alerted.

(6) Email in transit can be considered as though the message had been written on the back of a picture postcard. It can be read by anyone with the software and the motivation to read it.

Secure email mitigates or effectively removes these risks by using cryptographic techniques as explained earlier, and allows both businesses and private individuals to send email with a high level of confidence.

SECURE WEB ACCESS

As with email, many businesses have appreciated the value of having a website, whether as part of the sales function, for customer support, or purely as a promotional tool. Many of these sites have attempted to provide restricted access to items of value by the use of password-protected areas of the website. Unfortunately, passwords used in such a manner are effectively sent in the clear, that is to say, without any form of protection. Anyone listening to the network traffic can eavesdrop on password exchanges and store them for later use. As with email, unprotected connections to websites are open to attacks involving eavesdropping (sniffing), masquerading (spoofing), modification of data in transit, etc. Using secure e-commerce techniques, connections to websites can be strongly authenticated and protected through the use of digital certificates and suitable protocols such as Secure Sockets Layer (SSL).

SECURE EXTRA-NET ACCESS TO MAINFRAMES AND OTHER LEGACY SYSTEMS

Using secure e-commerce technology, systems can be deployed that allow a selected group of people or companies to gain access to certain resources within the company that would not otherwise be made available on the network. Such access might be provided via a website acting as a secure gateway to, for instance, a corporate database containing customer or product data. By being able to provide such access to selected parties such as important customers, suppliers, partners, etc., the company has created a stronger and more valuable relationship with them. They have effectively been given partial access to the inner sanctum of the company so that certain aspects of business can be conducted more efficiently and cheaply, or that new elements of business are now enabled. These extremely close working relationships with individuals or businesses outside our own companies can be enabled with great effect and a tight control on security, by using public key infrastructures as an underpinning.

VIRTUAL PRIVATE NETWORKS

Where a company's network is not connected to the internet, for example, and only authorised company employees have access to resources on that network, this might be called the company's private network. In such an environment, resources may be deployed and activities allowed that would certainly not be allowed if the company network were to be connected to the internet. Using new secure and standard protocols such as IPSEC, which has been developed in the Internet Engineering Task Force (IETF), it becomes possible to secure connections from one person or application to another, regardless of the type of network that connects them together. Thus, whether directly connected to the company private network or via the internet from halfway around the world, that network connection can, to all intents and purposes, be considered private to the company. The connection is protected from public scrutiny by using digital certificates and cryptographic techniques explained earlier. In essence, connections established using such techniques can be considered to be on the private network, or as it is commonly described, connected via a virtual private network. As an example of how this might be applied: a salesman in a hotel room in some foreign country can connect to the internet by dialling the local access number of his internet service provider. If he then sets up a virtual private network connection to head office over the internet, (using IPSEC for example), he can work just as if he were actually in the office.

ACCESS TO COMPUTERS

It is typically the case that in many large organisations, employees need access to several computers in order to be able to carry out their jobs. In some cases, access may be required on an ad hoc basis to say, thirty different computers, (databases, file-servers, accounts systems, etc.). Each of these computers may be password protected, so the poor employee has to memorise and manage up to thirty usernames and passwords.

SECURE PAYMENTS

Most people would understand that payment information (credit card numbers, account transfer information with authorisation codes, etc.), should not be exchanged without an appropriate level of security being involved. Secure e-commerce techniques have been developed that enable financial information to be exchanged in a safe manner. Many websites are now supporting the exchange of credit card information by protecting them with secure protocols such as SSL.

However the credit card companies themselves have been working on a credit card specific protocol to enable widespread and secure purchasing over the internet (SET – Secure Electronic Transaction). Due to the difficulty of building and managing completely interoperable deployments however, this protocol is not being taken up as quickly as the designers had hoped. On a related note, various other secure protocols have been designed to cater for sector-specific requirements or to integrate payments as just one phase of a much larger 'buying process'. These include such proposals as FIX (Financial Information eXchange), OBI (Open Buying Initiative), BIPS (Banking Internet Payment System), OFX (Open Financial eXchange), OTP (Open Trading Protocol), and others.

TECHNOLOGY AND APPLICATIONS

Any technology is only as useful as the applications to which it can be put. Secure e-commerce technology is no exception, so we must be able to demonstrate improved ways of doing things or enable entirely new and useful things to be done that could not have been accomplished before.

SINGLE SIGN-ON

It is typically the case that in many large organisations, employees need access to several computers in order to be able to carry out their jobs. In some cases, access may be required on an ad hoc basis to say, thirty different computers, (databases, file-servers, accounts systems, etc.). Each of these computers may be password protected, so the poor employee has to memorise and manage up to thirty usernames and passwords. This is extremely difficult and in many cases, passwords have been found written on sticky notes, attached to employees' screens. Any security offered by the passwords has just been rendered useless. One might suggest that merely setting all of an employee's passwords to be the same would solve this problem. However, in practice this is usually not possible because each system to be used may have different format requirements for usernames and passwords, and different criteria for password ageing and update. A further problem is that typically in organisations, an employee's job description will change from time to time, or employees leave, get hired, get promoted, or get transferred. All of this means that access to a different set of computers is required and user accounts need to be updated. Organisations can spend large amounts of money and time on merely managing this constant change. If some way were to be provided such that employees only had to sign-on once, to the company, and that after that, resources were to be made available according to assigned rights, the password management problem could be significantly reduced. Some products are beginning to appear on the market that now take advantage of the unique and secure identity offered by digital certificates, to provide such a single sign-on capability. This digital identity can be made portable and able to be carried around with the employee, in a personal security environment (such as a smartcard or software wallet) as described earlier.

AGENTS AND DOWNLOADABLE CODE

The download of unknown software has been an issue for some time. Most people will have heard of computer viruses and the havoc they can wreak on a company. A virus, however, is just one class of a range of malicious software that can do harm to your company. Furthermore, it is not just malicious software that should be guarded against: the use of untested unknown software (as can be downloaded from millions of sites on the internet) can cause just as much damage through faults (bugs) in the software. In some cases, you may not even know you have downloaded some code. Many websites, for instance, automatically cause the download of Java or ActiveX software to your computer, or plug-ins to handle the newest multimedia format. Software programs can be viewed merely as data, just as an email message or document. Consequently, it is possible to apply a digital signature to software programs, as was explained earlier with respect to email messages and arbitrary data files. Digital signatures can be used to provide authentication of the source of the software, and to show that the software has not been tampered with, since being issued by its author. With such protection in place, policies can be set up and enforced with regards to what software and from which sources, downloads will be allowed.

New secure e-commerce facilities are being discussed and developed all the time. For example, there is some interest in the development of trusted on-line negotiation facilities. The idea is to support the notion of several parties collaborating over the network to negotiate mutual agreement. Naturally, authentication of identity, and protection from repudiation will be important in such an environment, as would confidentiality and integrity protection.

Robert Bond

Head of Innovation & Technology Group, Hobson Audley

Robert Bond is author or editor of several books and journals on e-commerce and IT. He is chairman of the ICC (UK) E-Commerce Committee and legal counsel to the ICC Electronic Commerce Project at its world headquarters in Paris. He is a Fellow of the Society of Advanced Legal Studies and Companion of the British Computer Society.