Computer Science Tripos part 2 — Security

Lecture 12 - Robustness and the Legal/Organisational Aspects

Ross J. Anderson

1 Robustness

The cryptographic protocol failures that we have seen are mostly fixed quite easily by making some contextual information explicit. For example, we can fix the problem with the public key Needham Schroder protocol in a number of ways, but perhaps the simplest is to include the participants' names in each message.

In fact, experience shows that at all levels in the security engineering process, failures occur when assumptions – whether about cryptographic properties or about security goals – are not made explicit enough. Another example is the bank that wrote encrypted PINs on customer cards, without pausing to think about whether this encryption operation should be used to bind the encrypted PIN with the account number in some way.

Yet another example arose from experience with court cases over disputed electronic banking transactions. In a number of these cases, the customer denied an ATM or other transaction, and the bank claimed that the customer must be at fault since its systems were secure. However, when the customer disputed this and asked for disclosure of the bank's security policies and methods, the bank usually abandoned the case – its systems had not been designed with a view to withstanding examination by hostile expert witnesses. This mistake would not have been made had the system designers been explicit about their goals: these were at least as much about providing evidenced as reducing risk, and an appreciation of this would have led to a different design.

Perhaps more heed should have been taken of a 1977 Oxford address by Tom Watson, then chief executive of IBM:

> 'One of the most important problems we face today, as techniques and systems become more and more pervasive, is the risk of missing that fine, human point that may well make the difference between success and failure, fair and unfair, right and wrong ... no IBM computer has an education in the humanities'

On this note, we shall turn to the management end of the business. How can we go about reducing the risk that computer and communications systems can pose to the effective operations of our clients?

2 Risk Analysis

The first thing to do is assess the risk. There are a number of firms selling methodologies for this: some come in the form of do-it-yourself PC software, while others are part of a package of consultancy services. In fact, one of them – called CRAMM – is the standard imposed by the UK government on its computing sites and facilities subcontractors. The basic idea with such systems is to prioritise security expenditure, while at the same time providing a financial case for it to senior management.

The most common technique is the calculation of the annual loss expectancy, or ALE, for each possible loss scenario. This is just the expected loss multiplied by the number of incidents expected in an average year. A typical ALE analysis for a bank's computer systems might consist of several hundred entries, including items such as:

Loss type	**Amount**	**Incidence**	**ALE**
SWIFT fraud	£50,000,000	.005	£250,000
ATM fraud (large)	£250,000	.2	£100,000
ATM fraud (small)	£20,000	.5	£10,000
Teller takes cash	£3,240	200	£648,000

You will note that accurate figures are likely to be available for common losses (such as 'teller takes cash' – banks expect to sack about 1% of their staff every year for disciplinary reasons). However, for the uncommon, high-risk losses such as a large international funds transfer fraud, the incidence figure is largely a metter of guesswork.

Insurance can be of some help in managing large but unlikely risks. However, the insurance business is not completely scientific – for example, for years the annual premium for bankers' blanket bond insurance, which covered both computer crime and employee disloyalty, was 0.5% of the sum insured. This represented pure profit for Lloyds, which wrote the policies; then there was a large claim, and now the premium has doubled to 1% per annum. In these circumstances assessing risk is so difficult that many businessmen will prefer to leave it to the insurer (who in turn sends his inspectors to check out the insured's computer system and suggest security enhancements). In any case, such policies may have an excess of between £50,000 and £10,000,000 per incident, and so they only remove a small number of large risks from the equation.

The alternative approach, which is mandatory in the US Department of Defence, is the threat tree. We start out from each undesirable outcome, and work backwards by writing down each possible immediate cause. We then work backwards by adding each precursor condition, and so on. The idea is that by working round the tree's leaves we can see each combination of technical attack, operational blunder, physical penetration and so on which would break security.

This amounts in effect to an attack manual for the system, and in the US DoD it is treated as such – it may be very highly classified, and only a few people may ever be cleared to see it. Nonetheless, it must exist, and if the system evaluators or accreditors can find any significant extra attacks, then they may fail the product.

3 Computer Law

The legal environment can be important in civilian computer security, and in the UK there are four main pieces of legislation of interest to the computer security practitioner.

3.1 The Civil Evidence Act of 1968

This Act (section 5) treats the admissibility in civil trials of records generated by computers. Prior to this time, there was concern that many computer records might be hearsay and thus inadmissible. The problem was tackled by extending the already existing doctrine of a hearsay exemption for business records such as bankers' books, and the Act places computer records within this exemption provided that the system was used regularly to store and process information for regular activities, that the same kind of information as that produced in court was supplied regularly, that it was derived in the ordinary course of business, and that the computer was operating properly. Proper operation can be certified by anyone involved in operating the system – not necessarily an expert.

3.2 The Police and Criminal Evidence Act of 1984

PACE, as lawyers refer to it, performs a similar function for evidence in criminal cases. There, it is incumbent on the prosecution to show that there were no reasonable grounds for belief that the information presented was inaccurate because of improper use of the system, and that, at all material times, the system was operating properly – or, if was not, then this did not affect the production of the record or the accuracy of its contents. Whether the machine was working properly or not may require an expert such as the computer systems manager to testify. There is a recent proposal by the Law Commission to do away with this requirement in order to cut costs.

3.3 Data Protection

The Data Protection Act of 1984 was brought in under EU (essentially German) pressure. It set up a Data Protection Registrar; data users who hold personal information on data subjects in such a form that it can be manipulated automatically under program control must register and use the data fairly. All other uses of personal data are illegal. Following a further EU directive, it was updated by the Data Protection Act of 1998, under which the Registrar was renamed the Data Protection Commissioner, with the power to initiate investigations and prosecutions (previously, she could only act after receiving a complaint). It also extends the protection of the Act to paper records (although there is a lengthy transitional period for this).

Data subjects have a right to see their records and can demand changes if the information is inaccurate. They can also sue for damages (including distress, which is not normally actionable), but this right is little used. For his part, the Registrar can serve enforcement notices and even de-register a user; again, this power is little used, and one of the most frequent criticisms of the Registrar is slack enforcement (levels of registration are very low). There are in any event many exemptions, such as National Security; crime prevention (you can see your criminal record but not the tittle-tattle which may be held); taxation; social work; professional privilege; payroll and accounts; domestic data processing; and information which must eventually be made public in any case under other legislation.

Abuses which have received publicity include credit reference agencies keeping credit files by address rather than by name, and employers forcing potential employees to reveal criminal records by making enquiries under the Act.

The new Act has brought in tougher controls on sensitive information such as medical records, data on racial or ethnic origin, political and religious opinions and criminal offences. What these mean in practice will probably only become clear after the regulatory regime has settled down and the first few court cases have been decided.

There are also international political aspects. On the face of it, the new Act prevents personal data being sent from Europe to countries like the USA which do not have comparable controls. This is a source of tension between the EU and Washington. Opinions differ: some hold that the proprieties can be satisfied if US firms processing European personal data sign a contract undertaking to abide by European law (this is the approach taken by Citibank when processing German bank accounts in America). Others hold that, once some cases have finally been decided by the European Court in 2004 or 2005, many current business practices will be found to be unlawful, and this will cause severe disruption. Of course, the growth of the net brings a whole host of jurisdictional issues.

3.4 The Computer Misuse Act of 1990

This Act was passed in response to well publicised hacking cases, particularly that of the Duke of Edinburgh's mailbox (R v Gold and Schifreen). Under section 1, it is an offence to secure unauthorised access — 'to cause a computer to perform any function with intent to secure access to any program or data held in any computer', provided that you knew at the time that this access was unauthorised; the penalty is up to 6 months jail and a fine of up to £5,000.

Section 2 covers cases where such unauthorised access was used with the intent to commit or facilitate a serious offence; this raises the stakes to 5 years and an unlimited fine. Section 3 covers unauthorised modification of a computer system with intent to impair its operation or reliability; the idea is to make virus writing illegal (the first case is currently underway).

The Act has very broad territorial scope – almost any link with the UK will do. This takes account of the fact that hackers may operate in country A through a machine in country B to manipulate a machine in country C and get goods or services in country D.

Needless to say, the act has not been without its strange consequences. One of these was the conviction in Scunthorpe of a software writer who put a timelock in his software to enforce payment of the licence fee – until then a common enough industry practice.

3.5 The Interception of Communications Act 1985

Under this Act, the interception of communications (as in telephone tapping) requires a warrant from the Home Secretary. This has led to some interesting debate recently, in that it appears that if your ISP is also a phone company (as Demon is) then the police need a ministerial warrant to compel them to hand over your email; otherwise they can demand it (as material evidence stored on a hard disk) with a search warrant. Many ISPs would rather cooperate than insist on warrants; but most of them are not registered under the Data Protection Act for information disclosure for the purposes of preventing crime. Police initiatives are aimed at smoothing out this problem, and thus making ISPs free to disclose, while ignoring the tougher problem of whether ISPs ought to disclose in the absence of a warrant. Users have common law rights of confidence, and there are particular objections – for example, whether procedures will respect the special provisions for access to privileged material such as email between solicitors and their clients. All this is a subject of current lobbying.

3.6 Export controls

Current export regulations cover the shipment of physical goods, and are applied to most high quality computer security products (including cryptographic equipment, MLS systems and Tempest shielded equipment). Intangible exports – such as putting cryptographic source code on a web site – are presently unrestricted here (though not in the USA). European governments appear to have come under some pressure from the USA to fall into line, and a recent White Paper proposed closing the 'loophole' of intangible export. This aroused strong opposition, including from the Committee of Vice Chancellors and Principals, as it would have caused serious problems for international research collaboration and even for access by foreign students to teaching facilities here in UK universities. The issue will probably continue to be a live one.

3.7 Escrow

Much the most controversial policy/legal aspect of computer security at the present time is key escrow. I am not going to write at length on the subject here (I have done so elsewhere). Briefly, many governments are anxious that the spread of cryptography will impair their national intelligence and law enforcement capabilities, and seek to ensure that they continue to have the means of defeating any cryptography that is widely used.

There have been many proposals, from the US Clipper chip with its embedded family key and device key which provided the device ID and the session key respectively to authorised interceptors via the LEAF, through the previous government's proposal that every organisation providing cryptographic services should need a licence whose grant would be conditional on the escrow of confidentiality keys, to the present government's proposed electronic commerce bill which is similar except in that instead of licensing being mandatory, it is proposed that people who used licensed providers will enjoy a presumption of legal validity for their digital signatures. This is strongly opposed by civil liberties groups, and also by industry as likely to hinder the uptake of electronic commerce while providing very little benefit to the police. This is another story that will run and run.

4 Organisational Aspects of Security

Most security breaches in the real world come from blunders. Rather than the skill of a capable motivated opponent, they arise from the stupidity of the victim's own system designers, developers, operators or managers. We already saw this in 'Why Cryptosystems Fail' (which was appended to the 1b notes).

The organisational factors are particularly dangerous. Being a security manager is usually a poor career move, since it means saying no to people all the time; indeed, the average tenure of computer security managers in the US government is only seven months. Outside consultants are not much help; most firms, when faced with a choice between a well known consulting firm and a competent one, choose the former. What is usually sought is not a list of all the security weaknesses, but some 'due diligence' to provide career insurance. These issues are touched on in the appended article; the upshot is that many firms will really only pay for the appearance of security.

A recurring theme is the correlation between quality and security. For example, it has been shown that investment in software quality will reduce the incidence of computer security problems, regardless of whether security was a target of the quality programme or not; and that the most effective quality measure from the security point of view is the code walk-through. It seems that the knowledge that one's output will be read and criticised by a third party has a salutary effect on the average programmer!

Such indirect measures apart, knowing what computer and communications security is appropriate in a particular application, and how to sell it, is the real heart of the matter, and ultimately comes down to a matter of judgment. This may be influenced by techniques such as a risk analysis, and by the requirements of the client's insurers; it will also depend on the requirements of regulators (such as the Bank of England in the financial sector), on any contractual obligations imposed by customers such as the Ministry of Defence, as well as on the established industry practice and on whether the chief executive has seen a TV programme which sensationalised computer crime and drove him to demand some countermeasures.

In fact, there are clearly discernible fashions in computer security; during the mid 1980's, hackers were the main worry, and firms selling dial-back modems did a roaring business. From the late 80's, viruses took over the corporate imagination, and antivirus software made some people rich. Recently, with all the fanfare about the 'Information Superhighway', the firewall has become the new star product. Amidst all this noise, the computer professional must track the signal – distrust claims of expertise, and strive to understand what the real threats are.

5 Problems

1. Draw the threat tree for an automatic teller machine system.
2. Estimate the ALEs for a prepayment electricity meter system, on the assumption that fifty distributors sell tickets through 5,000 outlets for 2.5 million meters with an average turnover of £10 per month.
3. Would evidence from a 'Clipper' phone intercept be admissible in an English court?
4. In comp.risks 20.12, it was reported that people who downloaded an update needed to make Win98 Y2K compliant were also given an update to the Java virtual machine in Internet Explorer, whether they asked for it or not. This appears to be an attempt to show good faith compliance with a US court order concerning the Java licence. Could it have further effects under UK law?
5. The motor trade has devised a system which coats the insides of car bodies with about 10^5 microdots, each with a serial number which is visible under ultraviolet light. The idea is that this will cut car theft losses, as it will be very difficult for a villain to remove all the microdots. The insurance industry has asked you to advise on any possible pitfalls in this apparently splendid idea.
6. If a company used a multilevel secure operating system based on the Bell LaPadula model, should this have any effect on the admissibility or evidential effectiveness of its computer records?

7. Suppose you are the official in the post office responsible for franking machines. These machines dispense a fixed amount of credit against a ‘magic number’, along the lines of prepayment electricity meters. Your minister has now ordered you to put both machine and token sales out to private tender. Describe how this would affect your threat model, and what might be done to control the additional risk.

8. How would you tell whether a security evaluation scheme was really adding value by reducing risk, rather than simply being used for due diligence purposes?

9. Muffett has described the ‘Cinderella’ attack – “Hack the network time service and wind forward the clock until after January 2000, so the firewall turns into a pumpkin.” How realistic is this? What security-related failures or interactions would you expect to see around the millennium? Will the S2G turnover be significantly different?