PHYSICAL PROTECTION OF CRYPTOGRAPHIC DEVICES

Andrew J. Clark

Computer Security Limited
31/32 High Street, Dorset Place
Brighton, East Sussex. BN2 1RP
United Kingdom

ABSTRACT

With the growth of user awareness for the need to protect sensitive computer data by cryptographic means, this paper explains the need to protect critical cryptographic variables (particularly keys, and in some cases algorithms) in a secure environment within cryptographic equipment, particularly those used in the area of high value funds transfer transactions.

Design principles are outlined, leading to the concept of tamper resistant and not tamper proof devices to protect key data, whether the data be retained within physically large devices or on small portable tokens.

Criteria for the detection of attempts to gain access to sensitive data rather than attack prevention are outlined, together with two types of attack scenario - invasive and non-invasive.

The risks of attack on cryptographic devices are surveyed and intruder attack objectives are outlined, together with some typical scenarios. The available counter-measures are discussed. Several discreet mechanisms are described.

Typical detection mechanisms and sensor systems are discussed plus the design trade-offs that must be made in implementation, in particular manufacturing and maintenance costs versus scope of attack protection.

Once an attack is detected, various data destruction mechanisms may be employed. The desirability of active data destruction by "intelligent" means is proposed, together with a discussion of alternative techniques with particular reference to the data storage device characteristics.

Some experiences of tamper resistant research and development highlight the potential manufacturing problems - particularly in respect of quality assurance, product fault analysis and life-testing.

D. Chaum and W.L. Price (Eds.): Advances in Cryptology - EUROCRYPT '87, LNCS 304, pp. 83-93, 1988.

The desirability of tamper resistant standards and independent assessment facilities is expressed, the applicability of such standards and large scale protection methods on intelligent tokens, in particular smart cards and personal authenticators, is discussed.

What Do We Mean by Physical Protection?

Physical protection as applied to cryptographic equipment does not necessitate locking devices within mechanical safes or enclosing their electronics within thick steel or concrete shields, i.e. making them tamper-proof. It does, however, involve using sound design practices to construct a system capable of attack detection by a comprehensive range of sensors, i.e. tamper resistant. Few documents are available covering tamper resistant requirements and standards, among them are the Code of Practice for Cryptographic Equipment Security[1] and the U.S. Federal Standard[2]. They both require the incorporation of tamper-resisting, tamper- indicating, tamper-detecting and tamper-responding physical security measures. The level of physical security suggested should be such that unauthorised attempts at access or use will either be unsuccessful or will have a high probability of being detected during or after the event. Additionally, the standards recommend that cryptographic equipment should be prominently situated in operation so that its condition (outward appearance, indicators, controls etc.) is easily visible to minimise the possibility of undetected penetration.

This paper discusses both the concepts and detail of tamper resistant design to meet these requirements.

Why Do We Need It?

General principles of commercial cryptographic system design suggest that the security of data should not depend on the secrecy of the encryption methods used (for example, the algorithm) but rather on the secrecy of the key data. The generation and distribution of all key variables within a tamper-resistant environment such that no human being has knowledge of the key data in plain text is a necessary safeguard.

Classical Key Management hierarchies as described by Davies and Price[3] illustrate the need for high integrity key storage and

protection of top level 'master' keys. Clearly if these master keys are compromised by an intruder without detection the security of the whole system is brought into doubt. Attacks designed to discover such keys might involve physical penetration of the device to gain access to the internal components in which the keys are stored. This is particularly significant where many devices contain the same key, as may be the case in the design of some point-of-sale systems or networks of automatic teller machines. ATM's are almost certainly sufficiently well protected in their physical environment to defy attacks aimed at key discovery. Point-of-sale terminals however can be found on the counters of shops, large and small, and generally do not include particularly strong physical protection, largely because of the question of cost. It is quite possible that some point-of-sale terminals may be stolen with a view to compromising the secret parameters by an intruder. The fact that a terminal may be made unusable by the act of gaining access is immaterial to the intruder if profit is possible as a result of the discovery of the secret parameters, possibly by attacking the transactions at other terminals. What is required of the protection system is that the secret parameters should be lost or destroyed before access to them is gained.

Attack Objectives

An intruder's attack objectives are many and varied, but perhaps three of the best recognised are to:

i) Read cryptographic keys in plaintext

ii) Force cryptographic keys to a known or predictable value

iii) Render the crypto system useless and force the operators to revert to insecure methods.

Attack Scenarios

In order to achieve the objectives outlined above, various means may be employed - three typical methods are:-

i) Remotely monitor the mains supply to the crypto devices and attempt to analyse any conducted RFI to determine the plaintext data.

ii) Using a directional antenna and RF generator, produce a large RFI field in the vicinity of the crypto device to force the internal random key generators to latch-up and produce predictable keys.

iii) Gain access to the operations area and substitute a look-alike for the crypto device and steal the real device for subsequent attack.

This last scenario is probably the most real threat, particularly if one considers potential theft of devices containing live keys during shipment to operational sites or return for maintenance or, as in the case of POS terminals, from insecure shop premises outside normal office hours. Bearing this in mind, let us consider the response.

Defence Strategy

Clearly, as already stated, the major criterion is to design defence mechanisms that provide a very high confidence of intruder detection either during or after the attack, and in doing so make the cost of mounting such an attack greater than the intruder's potential gains.

Naturally, since different systems have different levels of potential gain, a "layered" approach to tamper resistance where varying levels of protection are available, is flexible and by its nature affords effective solutions in a cost conscious environment.

Defence mechanisms are broadly split into three main areas:-

i) Access control - for normal operator activity to ensure that unauthorised personnel may not operate the equipment.

(ii) Physical protection - generally mechanical, to ensure that casual substitution and non-invasive attack is difficult.

(iii) Electronics - mechanisms to detect intrusion - both invasive and non-invasive.

Let us address these areas in turn:-

Meeting the Challenge

a) Access Control

Since operational requirements will dictate that from time to time audit logs of cryptographic module behaviour must be dumped, master key updates initiated etc., operator access to cryptographic devices must be carefully controlled. This is generally achieved by the fitting of physical keylocks of the appropriate standard[2] requiring two or more trusted keyholders to initiate the top level security commands. For more stringent system operational requirements, these physical keylocks may be replaced with intelligent token interfaces, for example for smart cards or personal authenticators.

b) Physical

Generally, cryptographic units should be designed to be unique in appearance to avoid the possibility of casual substitution by a lookalike device. Physically strong mounting methods may be provided, although no direct access should be provided to the inside of the crypto unit. Ideally, there should be no ventilation holes, although if these are unavoidable they should be so constructed that it is impossible to gain access to sensitive areas within the device.

c) Electronics

In the majority of cases keys and other sensitive data are stored in Random Access Memory (RAM) with power supplied by independent battery sources, physically located close to the sensitive electronic devices. Alarm circuits are provided to detect intrusion and cause destruction of secret data. Some typical detector mechanisms are described in the next Section.

d) Detectors

(i) Dismantling

A wide range of sensors is available including simple micro-switches to detect removal of external case screws or lid assemblies, these may be supplemented by magnetic reed switches and permanent magnet actuators on mating surfaces. Active techniques of ultrasonic or infra-red space signature may be utilised, although because of power constraints it may be necessary to pulse these

detectors to conserve battery power. After an extended period on battery power, performance of these detection circuits may degrade and it is difficult to make them fail 'safe'.

(ii) Mains Power Variation/Monitoring

In order to ensure that no vestigial signal representing secret data appears on the mains power interface to the device, filtering should be employed between the device mains input and the power supply input point, and the power supply low voltage outputs should be adequately filtered and decoupled. Passive transorbs and fuses provide protection against deliberate over-voltage and reverse-voltage attacks on the device while good design practices must be observed when implementing power up/down monitoring circuits designed to protect the integrity of secure data.

(iii) Physical Removal

Unauthorised attempts at moving the device can be detected by tilt and jitter sensors which operate when the device is, for example, tilted more than 20^{o} from the horizontal or subjected to the sort of vibrations generated by a normal power tool. Additionally, to protect against illegal removal of the power or communications cables, closed-loop alarms should be connected through both security devices and peripherals via the connecting cable assemblies.

(iv) Drilling and Grinding

Encapsulation of the sensitive electronic components holding secure data in a potting resin is a well-known process which certainly acts as a good physical barrier to an intruder wishing to probe the key storage electronics. The simplest method to gain access to the sensitive components is to drill, mill, grind or plane the potted area until sufficiently close to the target and then proceed more carefully using fine hand tools. In order to successfully attack in this way, knowledge of the layout of the PCB and the associated components is desirable and this is best accomplished using X-Rays, the drilling procedure may then be undertaken more accurately. Embedding a fine mesh of multiple layers of randomly located fine wires within the potting or, alternately, integrating a flexible PCB with multiple orientation alarm tracks on it, is a useful detection mechanism against these attacks. It is interesting to note that if the wires are fine enough, accurate detection of their location by

X-Ray means is a relatively difficult task. Obviously, all components accessing secure data paths must be enclosed within this encapsulation. In a classical bus-oriented micro computer solution, this obviously applies to all devices having access to the main data and address busses.

(v) Solvents

Since the embedded shield methods of Section (iv) render drilling, grinding and planing relatively difficult, a suitable chemical solvent attack on the encapsulation would prospectively seem attractive. If the potting compound has been carefully selected by the manufacturer such that any appropriate solvents for it are also solvents of the chip fabrication materials and PCB fabric, together with probably having special handling problems owing to its volatility, then solvent attack becomes more difficult. Embedding fusible links within the potted area such that mass flooding or immersion is impractical is an added safeguard.

(vi) Temperature

Since the majority of electronic components perform within a temperature specification of, typically, -30°C to +85°C and these would generally include the alarm detection and key destruction circuitry, rendering these circuits inactive by raising or, more generally accepted, lowering the unit temperature to, typically, -80°C would render these circuits inactive. Hot and cold temperature attacks are relatively easily detected by the inclusion of temperature sensors within the alarm circuitry which operate at, say, -25°C and +70°C although the effect of thermal shock on these devices and the units themselves, due to sudden change in temperature (e.g. by immersion in liquid nitrogen (-195°C)), must be carefully calculated to ensure correct failsafe operation. The choice of temperature detection thresholds is important if false alarming of a device in transit (e.g. an aircraft hold or a car boot) is to be avoided.

(vii) X-Ray

As mentioned in Section (iv), the use of X-Rays as a mechanism for locating critical components and data paths is extremely useful. Including X-Ray detection in alarm circuitry at first sight seems attractive, although in practice when devices are despatched from

manufacturers premises for subsequent shipment by air freight, they are likely to be X-Rayed under normal security procedures and hence alarm systems activated. As an alternative the sensitive component areas may be screened against X-Ray surveillance by a lead shield. Practical experience shows that, to be effective, the thickness of lead should not be less than typically 3mm, and the surfaces should be stippled and scratched in random pattens to enhance the deflection effects.

(viii) EMI/RFI

In considering the effects of electromagnetic and radio frequency interference, it is apparent that these effects are bi-directional i.e. radiation of signals from the device should not be capable of interpretation to reveal secret data, nor should any external interference source directed at the unit cause it to malfunction or 'latch-up' into a predictable state. This latter effect is particularly important in considering the behaviour of white noise seeded random number circuits which generate encryption keys. In designing device enclosures, material choice and bonding techniques which affect EMI behaviour are naturally important. It is generally accepted that metal case construction is preferable, and good electro/mechanical designs should be employed to ensure minimum escape of radiated energy. Additional barriers around sensitive component areas may be provided using copper screening cans with modular 'onion skin' construction techniques. Recent advances in spray-on conductive graphite, nickel and silver coatings give EMI/RFI attenuation performance figures of typically greater than 70dB which approach good design objectives of, for example, 100dB. A combination of these spray techniques and metal case construction can lead to good EMI/RFI resilience and a reasonable level of physical strength.

Data Destruction

It is generally accepted that all sensitive data storage requires a "zeroisation capability".[1] This implies that all data bits are actively set to zero following an alarm condition. Although at first this appears a simple task which may be accomplished by disconnecting the power supply of the volatile storage devices or loading a single set of zero's into non-volatile storage, in practice this is unsatisfactory. Certainly random access memory (RAM) devices have

residual data retention characteristics which, under controlled conditions, permit reconstitution of the original data contents prior to the attempt at erasure.

In order to be confident of destroying the data successfully, all storage cells must be actively purged by overwriting with all '1's, and then all '0's at least three times in rapid succession, followed by the shorting of the device power supply input pins to ground. In cases of extreme sensitivity it is possible that the only acceptable method of destroying the data is by non-reversible physical destruction of the storage devices themselves, although this is naturally rather a difficult strategy to follow in a production test environment!

Tradeoffs

In any system containing detection and destruction methods as described here, there is naturally a cost penalty for providing very high levels of tamper resistance, due to construction and test requirements by the manufacturer. It is naturally important to analyse the risks of key disclosure against cost of protection and specify a suitable implementation. For some of the methods described here, where the tamper resistance cannot easily be removed for maintenance purposes, the implications of a throw away replacement maintenance policy should not be overlooked.

Intelligent Tokens

The attack scenarios and associated countermeasures discussed so far have been oriented towards physically 'large' devices. With the growth in EFTPOS schemes and the increased use of smart cards and personal authenticators, the applicability of these techniques is worth considering.

Since intelligent tokens are 'stand alone' devices prone to theft and/or substitution, good cryptographic system designs should ensure that compromise of one token should not threaten the security of the whole system. Generally these units contain user related key data stored in battery backed-up RAM and mask programmed ROM. Naturally, size constraints dictate that protection against X-Ray inspection are impractical and hence mechanical attacks using precision micro-manipulation techniques are the biggest threat. Recent advances in

token design have taken this into account, and the sensitive data areas on the device silicon structure have been 'buried' under several layers of metallisation and convoluted data paths constructed to add confusion. These simple techniques are probably the best that can be practically applied in volume manufacture and hence the risk analysis of token-related key compromise must be stringently assessed.

Independent Assessment

Few independent assessment houses or laboratories seem to address the rather particular needs of tamper resistance testing. The author has had experience of two such organisations, they are:

The National Physical Laboratory,
Teddington, Middlesex, UK,
and
TNO Division of Technology for Society,
Delft, The Netherlands.

The role of such bodies is to act as an independent test and assessment facility for organisations specifying and procuring items of cryptographic equipment with tamper resistant capabilities.

Future Trends

As markets develop and grow within the financial sector the need for secure product design to keep up is unquestionable. The potentially explosive growth of EFT/POS schemes has led to research in the area of low-cost tamper resistant modules where either the unit cost is so low that the secure components can be thrown away if they fail, or the tamper resistant mechanisms are reusable allowing return to factory maintenance.

The level of protection afforded and complexity of attacks catered for must naturally increase with time - particularly since current trends in computer system design are moving more into a distributed architecture approach which makes physical computer protection, particularly personal computers, much more important.

Some manufacturers have already looked to miniaturise their crypto key storage components, and the inclusion of data destruction circuitry on-

chip is an integrated part of this process. To date these techniques have primarily been non-recoverable, i.e. they physically destroy the silicon - adaptive systems will no doubt appear in due course.

Underlying all of these trends however must be the message that as systems designers we must continue our R & D programmes to devise new techniques and strategies to meet the ever increasing sophistication of computer crime.

References

(1) "Cryptographic Equipment Security: A Code of Practice" S.C. Serpell, British Telecom Research Laboratories, Martlesham Heath, Ipswich.
Published by Institution of Electronic and Radio Engineers, 99 Gower Street, London WC1E 6AZ

(2) "U.S. Federal Standard 1027 Telecommunications: general security requirements for equipment using the data encryption standard"

(3) "Security for Computer Networks - An Introduction to Data Security in Teleprocessing and Electronic Funds Transfer". D.W. Davies & W.L. Price
Published by John Wiley & Sons