# Fault Induction Attacks, Tamper Resistance, and Hostile Reverse Engineering in Perspective

David P. Maher, AT&T Labs, dpm@research.att.com

Abstract

We put many of the new fault induction and reverse engineering attacks on secure systems into the context of real device implementations and actual systems. We describe countermeasures that diminish the overall practical significance of these new results when considered in the context of a rational design process and an overall systems security strategy.

## 1. Introduction

For years, secure systems engineers have used a number of technologies to resist reverse engineering attacks on their systems. Although reverse engineering is probably ancient, it has until recently been a rather obscure art. Secure systems designers have known for decades that an attacker can intentionally induce faults in the operation of a system thereby exposing cryptographic keys or bypassing security functions. Although the design of countermeasures to malicious fault induction has not been a popular topic at conferences, it is an art that has been carefully practiced and which includes well-structured processes. It is also an art that draws (but not exclusively) upon the far less obscure science of fault-tolerant systems.

Recently, a number of research notes have been published over the Internet regarding fault induction attacks and reverse engineering (overcoming Tamper resistance) [BDL96], [AK96a, b, c] [BS96a, b]. These notes have received considerable press attention, some of it stimulated by the researchers involved or their organizations. For example, Bellcore published a “Security Alert” later followed by a press advisory to trumpet the results of certain discoveries made by their scientists in the area of fault analysis. In the Security Alert and the Press Advisory [BEL96a,b], Bellcore claims their results will “cause dramatic changes in the marketplace” and that fault induction is “such a novel approach to breaking cryptographic security systems that it is considered a new threat model.” Bellcore further says: “of all the challenges facing electronic commerce -- billing logistics, Internet congestion, lack of privacy, and so forth -- the new attack on tamperproof devices may be the most debilitating.“ Ross Anderson has published comments on the Internet regarding his papers with Markus Kuhn, and according to the London Telegraph: “Anderson said his latest research indicated that two of the world's most widely used systems for encoding sensitive financial information - the RSA and DES encryption standards used by most banks - could also be cracked easily." The Telegraph concludes from their interview with

Anderson that: “their discovery could spell the end of the Mondex system, which relies entirely on the security of the smartcards for its integrity. ‘I don't think you will be able to have floating systems like Mondex any more, where all the information is held on the smartcards,’ said Anderson.”[LT96]

These statements are quite alarming. Well over $100 Million has been invested in the development of Mondex, and new investment by 17 Global Founding Members and by MasterCard International has raised expectations that a global electronic cash system may be possible. Smart cards are thought to be the key to the security necessary for a number of very important applications. The statements above, while raising the consciousness of some people, are generating a lot of very unproductive discussion and ill will. This paper provides some perspective on the recent discoveries, places them in historical context, samples some of the effective countermeasures to some of the attacks, and comments on their practical implications in the context of the security design practices of professional secure system designers.

Ross Anderson and the Bellcore authors have not broken DES or RSA and they do not appear to be on the verge of doing so. They *have* shown that they can break poorly designed systems that use DES and RSA and well-designed systems that are misapplied. The thesis in this paper is: by using a good system design process and risk management procedures, and by proper application of threat models, it has been and will continue to be possible to design systems with the security properties that are necessary to ensure the success of a given product. Once we fully understand what is new about the recent research announcements as well as what is old, and once we understand the design approach taken by seasoned secure system designers, the reader may be less inclined to prematurely accept declarations regarding the demise of certain systems. Indeed the reader may be inclined to judge that there are reasonably good prospects that security for electronic cash systems such as Mondex can be sufficiently effective.

## 2. Fault Induction Attacks Are Old

In [BDL96], Boneh, DeMillo, and Lipton from Bellcore describe ways in which a secret cryptographic key can be mathematically derived from results of certain cryptographic calculations in cases where a microprocessor can be induced to make errors during the calculations. Bellcore, in the press hype published on the Internet, claims that the very idea of attacking cryptosystems by inducing them to make errors is novel (“a new threat model”)[1]. They gave it a name: “Cryptanalysis in the presence of hardware faults.” The Bellcore researchers suggest that a cryptographic device can be induced to make failures by subjecting the device to environmental conditions

[1] The Bellcore Press Release response to the question “What is your new threat model?” is “We observed that once a computing device performs a faulty computation, it might leak information that can be useful for inferring secret data. This is a novel approach to the widely acknowledged fact that no computing system is safe from faults. “

outside the expected range of operation. They suggest subjecting the device to temperature or radiation extremes, as an example. However, this is an old threat model that is certainly well known by anyone who has designed military grade crypto gear. Indeed, over a dozen years ago, I began work on a series of documents, appropriately called Security Fault Analysis (SFA) reports, that analyzed failure modes of cryptographic hardware. In these reports I was required to note the cryptanalytic consequences of many different classes of failures down to the gate and transistor level. I used a standard and formal methodology for doing this. The overall methodology was extended to firmware used in security devices. Over the past decade or so, as many crypto device designers have focused on commercial applications, the threat model has become well known among the community of commercial security experts, including many smart card and encryption chip designers. In fact, attention to the threat model is demanded by a commercial grade standard known as FIPS 140-1 [FIPS] which was published in final form in 1994, but in draft form several years earlier. To wit:

> "Electronic devices and circuitry are designed to operate within a particular range of environmental conditions. If the devices or circuitry are operated outside of this range, their correct operation is not guaranteed. *Deliberate or accidental excursions outside the specified normal operating range can cause erratic operation or failure of the electronic devices or circuitry within a cryptographic module that can compromise the security of the module***.** In order to provide reasonable assurance that the security of a cryptographic module cannot be compromised by environmental conditions, the module may either employ environmental failure protection (EFP) features or undergo environmental failure testing (EFT)."

The FIPS standard specifically anticipates the possibility that failures may allow secrets to be revealed:

> "The protection features shall involve additional electronic circuitry or devices that shall *continuously* measure these environmental conditions. If a condition is determined to be outside of the module's normal operating range, the protection circuitry shall either (1) shutdown the module to prevent it from operating outside the normal range, or (2) *immediately zeroize all plaintext cryptographic keys and other unprotected critical security parameters*. Documentation shall provide a complete specification and description of the environmental failure protection features employed within a module."

Fault induction is anticipated by the ISO standards for smart cards, where designers are admonished to detect variations in clock frequency and voltage, for example. Thus, it is difficult to understand how anyone can credibly claim that this is a new threat model. We have been dealing with this model for many years as engineers and

scientists from companies such as Motorola, Siemens, Intel, Hitachi, AT&T and several others can attest.

Over the years a number of countermeasures have been developed to deal with this "new" threat model. They include memory access control, detectors for variations or out of range conditions for voltage, clock frequency, light, and temperature, and detectors for removal of a chip's passivation layer. The quality of the implementation of the countermeasures and even the decision to include any of them in a given security device is usually dependent on an overall threat analysis for the applications for which the security device is designed. Until recently, it did not pay to include these countermeasures in the vast majority of chips that carry out security functions. In chips and other systems where they are included, they are typically not the only defense against fault induction attacks. On the other hand, just as there has been hype regarding attacks, there is probably more hype regarding the security features of chips that defend against intruder attacks.

So the Bellcore threat model is not new, but there are some novel and interesting results in the paper. The Public Relations spiels seem to imply that these new results are "dramatic" and the problems they cause are most "debilitating" for electronic commerce. We'll see that these are exaggerations, and that many electronic commerce applications already provide multiple defenses, and most other applications will be able to defend against the attacks without great pain.

More recently, Anderson and Kuhn published on the Internet some improvements on the Bellcore attacks. We'll see that a well-designed system can counter these improvements as well. Their note describes attacks on device memory. These attacks are not new (though it is not clear that they are claimed to be). Similar attacks are even easier than described in their note. For example, "One time programmable ROM (OTPROM)" is usually just an erasable and programmable memory (EPROM) in a different (cheaper) package. Such memories are easily modified, and this fact can be used to attack the security of a system implemented with code stored in such memories.

Secure system designers have typically included alarms on memory integrity. A cyclic redundancy check over the memory spaces is usually sufficient, but some systems can go so far as to explicitly check the integrity of the alarm itself. When the CRC detects a fault, the system resets and a failure counter is incremented. Once a threshold is exceeded, the system can purge sensitive cryptovariables. These measures are usually sufficient to thwart attacks on ROM, EPROM, and EEPROM, although designers are folly to rely solely on them. Other measures such as sanity timers, function time-outs, and the sprinkling of reset commands in the program memory are also commonly used. Some applications go even further and use cryptographic checksums and application authenticity checks that are much stronger than CRCs. Just because a chip doesn't implement specific countermeasures such as those mentioned here does not mean that the chip cannot be used to implement security functions. As explained

below, the specific countermeasures and how they are combined will depend on the application and the threat model.

Anderson and Kuhn's paper [AK96a] on Tamper resistance includes little that has not been known to many in the secure systems design community, with the exception of some details of Kuhn's successful attack on the Dallas Semiconductor chip. However, it is nonetheless a well-written and useful survey of attacks on chips that claim to be "Tamper-proof." In contrast to the Bellcore paper, actual attacks on chips are described in this paper. It teaches to all the lesson that good designers already know from experience, test results, analysis, and (one might argue) common sense: Nothing is Tamperproof – anything that can be built can be destroyed.

## 3. What is New?

Boneh, Demillo, and Lipton's results are relevant to a very broad range of authentication, encryption, and digital signature protocols. They describe new attacks whereby secret keys can be discovered when faults are induced in the computations involved in digital signatures and authentication protocols *under certain conditions*. With the conditions properly stated, the attacks are theoretically sound and typically require a cryptographic function to be applied to a datum correctly and on the same datum at least once incorrectly. For example, the Bellcore authors show that under certain conditions, if a smart card is used to digitally sign a message using a secret key that is supposed to remain concealed within the card, and if the message is signed once correctly and the same message is signed at least once incorrectly (say, because an intruder has caused the microprocessor to miscalculate during the computation of the signature), then an intruder can extract the secret from the results of the two signatures. They give very simple formulae for the computation of the secret. These results appear to be new and they are of great interest.

After hearing of the existence of the Bellcore fault induction attacks, Arjen Lenstra, of Citibank, independently formulated a new attack on some implementations of RSA that use the Chinese Remainder Theorem. Lenstra's attack appears to be stronger, as it only requires one faulty signature and does not require a correct signature. With the CRT, a signature is computed twice, once modulo each of the prime factors of an RSA modulus n=pq. Then the CRT is used to compute the signature mod n. Lenstra observes that if one has a message x that is to be RSA-signed using a secret exponent r, if an intruder induces a fault during the calculation of precisely one of $x^r$ mod p or $x^r$ mod q producing a faulty signature S mod n, then either p or q is, with extremely high probability, equal to gcd($S^v$-x, n) where v is the verification exponent. In contrast to the Bellcore press claims, Lenstra has been precise and careful concerning implications of his result.

Biham and Shamir [BS96a] describe new attacks on hypothetical implementations of DES that use fault induction. An older version of Mondex uses DES, but is not vulnerable to this attack for much the same reasons that the new version is not

vulnerable to the Bellcore and Lenstra attacks (see below). The only fault mechanisms that appear to have a good chance of working here appear to be those described by Anderson and Kuhn in [AK96c]. In fact, they improve significantly on the Biham Shamir attacks in efficiency. However, the AK attacks are not practical, simply because they are typically defeated by a number of common practices, such as memory integrity checks and newer approaches that include application signing, mentioned above. Most if not all of communications security and Electronic Commerce gear that I am familiar with have built-in defenses to these attacks. It appears to be worthwhile to contemplate what kinds of integrity checks on memory are necessary in order to be sure to thwart a determined attacker. If one assumes a powerful opponent with the ability to change memory at will, some of the weaker integrity check mechanisms may not be sufficient.

In principle, dynamic attacks on program memory can bypass memory integrity check processes. Thus, other defenses may be required when called for by the threat model.

## 4. Countermeasures to the New Attacks

If the threat model for a given application indicates that resistance to fault induction attacks is required, good designers will determine the kind of access that an intruder may have to the device and then ensure that their systems are properly resistant to these attacks. Designers should consider the possibility that faults will somehow be easy to induce (even if they appear to be difficult). In other words, there need to be multiple layers of defense against this type of attack. There may be multiple layers of defense in existing systems even though the specific attacks were not anticipated, as is illustrated below. This is because although the Bellcore attacks and their improvements by others are new, they belong to general classes of fault induction attacks that have already been analyzed. Let's go over some countermeasures:

1. The Bellcore attacks and their improvements assume that faults can be successfully induced. There exist a host of countermeasures against various methods of fault induction, such as the various detectors mentioned above. Radiation, though mentioned by the Bellcore press releases as an effective fault inducer, is very unlikely to work on today's CMOS chip designs. Chips are more likely to stop functioning altogether than to fail in ways that permit the Bellcore attacks to be carried out. The rumors of successful attacks using microwave ovens are belied by the fact that kitchen microwaves operate at 2.45 GHz – that's a wave length of 15 cm. Features of smart card chips are now submicron. Highly focused radiation (such as a focused ion beam) could work. Attacks on memory as suggested in [AK96c] are much more likely to succeed, but we have already discussed above the various alarms, alarm checks, and other countermeasures used against memory fault induction. Nonetheless, we will assume that these physical and logical defenses against fault induction will fail. The author is aware of other attacks, (not yet published) that can induce faults (those

security fault analysis reports were not produced for nothing). We can, and often should assume that failures will be successfully induced. Then what?

2. The Bellcore attacks assume that the same message will be signed more than once. Many security protocols and system design requirements forbid this. If they don't, there is usually no reason why a device needs to sign precisely the same message twice. One can usually add a sequence number or, better yet, random noise to some field in each message that is signed making the message unique. This will thwart the Bellcore attacks.

3. Making the message unique will not thwart Lenstra's attack mentioned above. However, if a random cryptographic value is added to the message x, and if the value is not revealed and not used again, then Lenstra's attack appears to be countered. This is because Lenstra's attack requires that the intruder knows the message x. If the signature, S, is computed incorrectly, then the intruder will not be able to recover x using the public key. If the message is signed correctly, no harm is done even if the message is correctly recovered by the intruder using the public key.

4. Another defense against the Bellcore attacks and its generalizations is to verify the signature before revealing it. In the case of the most commonly used signature method, RSA, one can verify the computation was performed without fault just by applying the public key signature verification function. Exponentiation complexity is typically a linear function of the length of the exponent. So if a short verification exponent such as binary 11 is properly used (see the attacks in [CFPR96] on short exponent encryption), the performance penalty is typically only 1% or so for good-sized moduli. The penalty may be greater, percentage-wise, if techniques for secret exponentiation such as addition chains are used, but this is often not the case with smart cards due to the limited availability of RAM.

5. The Bellcore attacks and Lenstra's attack on the RSA implementation assume that the intruder will know the modulus. Thus, an effective defense is to protect the system parameters such as the specific modulus used in a signature calculation and its verification. This can be done with closed systems such as might be used for ecash protocols implemented on smart cards. There is often no need to make the modulus public, and various cryptographic means can be used to protect a specific modulus in key exchange protocols.

## 5. Further Considerations

What is the probability of success of a given attack and how much work is required to carry it out? The Bellcore security alert says the very possibility of an attack's existence is a "sign of great danger." This is absurd. As a secure systems designer, if I had to consider every attack as dangerous and address every conceivable threat, I'd end up grossly overdesigning every system. It is much more reasonable to analyze the threat in the context of the application using a threat and risk analysis methodology.

One cannot categorically say that a given attack is dangerous without understanding the cost of the attack as well as the payoff. Anderson and Kuhn help us make some headway with this question in the area of reverse engineering, but the examples in the survey [AK96a] are still very general. Layered countermeasures can interact. Obscurity can increase the work factor. Attacks that are simple for one chip may be very hard for another.

Payoff for an attack is heavily dependent on system design. One can increase the effectiveness of a system by decreasing the payoff on successful attacks rather than just making the cost of success high. Knowing the expected cost of the attack and the expected payoff will help to determine whether one should expend any effort on further countermeasures. It seems that pirates have mercilessly attacked the NewsDatacom satellite video encryption system by very straightforward reverse engineering. However, the launch of periodic countermeasures by the designers, together with legal measures taken against organized attackers increase the risk and limit the payoff keeping overall losses suffered by the DirectTV system relatively low. The pirates succeed in getting free service, but it is poor service: "Mommy, mommy, we can't get the cartoon channel any more." "Sorry kid, we'll have to wait a few weeks until we can download a new solution to the latest Electronic Countermeasure." This scenario is not realistic in most households. From the point of view of the attackee, an attack is successful only if the attackee feels the sting. It is not clear at all that these satellite pirates have had that kind of success. The renewable security strategy of the satellite systems providers appears to be mainly effective in spite of what appears to be very poor system design and inherent weaknesses at the point of the smart card interface [McC94].[2] On the other hand, this satellite video industry is training an army of smart card hackers who can turn their attention to electronic cash systems as soon as they see sufficient incentive.

The Mondex system controls payoff through the use of "purse classes." You cannot buy a Ferrari or anything of large value with a consumer Mondex card. Merchant cards that hold high value are subject to more controls than consumer cards. On the other hand, consumer cards interact with bankcards, and certain aspects of the card behavior are thereby accounted. Mondex has partial accountability and can be fully audited. A number of mechanisms cooperate to make it more difficult for an intruder to successfully exploit a successful reverse engineering or tamper attack.

Another relevant question in this context concerns whether there are effective recovery techniques in the event that an attack is successful. For example, if a fraudster attacks successfully and creates some electronic cash illicitly, can I detect the event somehow and shut down the fraudster before too much damage is done?

[2] Its interesting to note that terrestrial video services providers (cable companies) that have implemented fixed security systems inside their set top boxes have indeed felt the sting. They are losing significant amounts of money and are looking for security that one executive characterizes as "bloody overkill."

Good high-security systems will anticipate the successful fraudster scenario, no matter how good the attack prevention is. Once Mondex detects that a Mondex card has been cloned, an authenticated virus may be invoked to limit consumer-to-consumer purse interactions, forcing the clone cards to interact with on-line purses in order to seek the payoff.

## 6. Discussion of Some Specific Systems

The Mondex electronic cash system was mentioned as being vulnerable in both the New York Times article [NYT96] on the Bellcore attacks and in the London Telegraph [LT96] article on the Anderson-Kuhn improvements. To see how appropriate the hype is, let's see how Mondex fares at least in design. The rollout version of Mondex uses public key digital signatures to ensure the integrity of messages in the Value Transfer Protocol. Mondex indeed has several layers of defense against the new attacks on signatures. Details have been omitted (see the discussion below on security by obscurity):

1. There are several physical and logical defenses against most failure induction attacks such as those hypothesized by the Bellcore people and others from years ago. These include defenses against attacks on memory such as hypothesized by Anderson and Kuhn.

2. Should these physical and logical defenses fail, there are protocol level defenses. The Mondex Value Transfer Protocol complies with the CEN standard 1546. No message is ever digitally signed twice. This defends against the Bellcore attacks (and the older version of Mondex is secure against the Biham-Shamir attacks).

3. Should the protocol itself fail to thwart an attack, there is a software security layer defense that will frustrate the Bellcore and the Lenstra attacks.

4. Should these fail, there is a powerful message layer defense that will also defeat the Bellcore and Lenstra attacks.

5. In the event that all of these independent defenses should fail, Mondex also limits the scope and utility of the secrets that are stored in the card, thus limiting the damage done. The payoff for an attack is contained by purse class limits and accountability requirements for higher class purses that store large value.

6. Since the Mondex security team believes that it is prudent to assume that some new completely unanticipated attack could arise that is not caught by the above countermeasures, Mondex also employs an active risk management system that allows Mondex to recover from compromises, and discover their source.

7. Finally, the active security system in each card can be changed, and the security system for all of Mondex can be upgraded over time, presenting a moving target for hackers. Mondex can use one signature scheme initially and then migrate to the use of another signature scheme, and then another. The design team took pains to make renewable security part of the overall security approach precisely because they believe in the Gospel that Ross Anderson and others have been teaching: Every security scheme will eventually be broken.

The reader may be annoyed that the details of the Mondex defenses have been omitted. The reader is invited to curse at the remarks below on "security by obscurity."

It is interesting and encouraging to note that the Mondex defenses were designed and implemented well before the Bellcore attacks were discovered. Mondex has designed the security of their system to withstand broad classes of intruder attacks that go far beyond the specialized attacks from Bellcore and their improvements. Since the security scheme is renewable, stronger defenses can be added over time. Of course, there are many other attacks on other aspects of Mondex. It is not possible to go over the entire threat model here. Indeed, the threat and risk analyses fill volumes, and they have been accumulated over more than six years.

Another system that is interesting to analyze is the security chip developed for AT&T's IVES (the Information Vending Encryption System). The chip is now sold by Lucent Technologies and VLSI Technologies, Inc. [HM95]. The IVES chip uses the same antifuse memory technology "VROM" used in the infamous Clipper and Capstone chips. IVES secrets are stored in this memory. Anderson and Kuhn, in [AK96a] catalog a variety of attacks on memory, but these attacks are unsuccessful against this memory technology. However, they do remark that they have it upon reliable authority that the Clipper Chip was "reverse engineered" by a US manufacturer. Looking beyond the annoyance of a publicly passed but unsubstantiated rumor in a research paper, and a lack of precision about what was actually produced by this reverse engineering and how "blind" it was, I know that it is possible for the VROM anti-fuses to be read through very tedious destructive methods (acknowledged by Anderson). It requires the destruction of many chips in order to recover any single secret of hundreds of bits that is common to the chips. However, the IVES chip was designed in such a way that the binary vectors in the VROM are uncorrelated from chip to chip. This is true even if system common secrets are stored in these memories. It appears to be extremely unlikely for an intruder to get a complete secret from just one chip by attacking the anti-fuse memory.

The IVES chip was designed to function in very hostile environments. Like the Dallas semiconductor chip successfully attacked in [AK96a], the IVES chip has an externally accessible bus that potentially could be used to attack the chip. However, the IVES chip shuts off the external bus when any security procedure is running.

Secrets never show up on the externally accessible bus. In fact, during security kernel operations, no instructions or data at all appear on the external bus. They only appear on the internal, protected bus. Thus, the chip cannot be reverse engineered by the methods used on the Dallas chip. Once again, there are certainly other attacks, but there are also other countermeasures against the attacks we know of as well as attacks that are yet to be discovered.

## 7. More Observations on Reverse Engineering and Tampering

It should be observed that the fact that the bus encryption capability of the Dallas chip was broken by the Kuhn attack does not itself mean that any given application of the chip is in jeopardy. Typical applications of the chip include controllers in PIN pads and ATMs. In these applications the intruder does not take the same kind of possession of the device that the owner of a smart card does. The Kuhn attack requires the intruder to access the bus for a good period of time to set up a test interface. This is unlikely in typical applications where the device is protected by locks or otherwise. It is not made clear in [AK96a] what the intruder will gain when attacking any of these applications especially since the secrets are session keys that are changed constantly. Perhaps there are some ways that an insider can use this attack. That should certainly be considered. Breaking the bus isolation feature of this chip is probably not the devastating event that some people have been led to believe. The attack has been known for several months and it does not appear to be necessary to modify much, if any of the most commonly used banking equipment. The chip does serve one of its intended purposes of ensuring the uselessness of information leaked from the bus onto other computer interfaces including coprocessor interfaces.

It is perhaps worthwhile here to dispel the rumor passed on to the London Telegraph that Mondex uses a Dallas DS5000 series chip, successfully attacked by Kuhn, as documented in [AK96a]. It is certainly not used, and it has not been considered for, and is not even appropriate for the Mondex secure purse application. It is conceivable that the rumor arose from the possibility that the chip might be used in some kind of interface device such as a wallet or Point-of-Sale terminal. However, the integrity of the Mondex protocols does not depend on the security of the interface devices.

## 8. In Defense of “Security by Obscurity”

Much of the rumor and frustration about systems like Mondex and IVES stems from the security policy of “Security by Obscurity,” a policy that is often deplored as stupid. At some risk, I will argue that such a policy can be part of a rational strategy to decrease overall system security risk. Obscurity may be more than adequate to counter some threats. Obscurity can increase the risk (affecting the business case) of some attackers. Organized crime has always been based on rational theories of investing. They may skirt the law, but they still try to make money the easiest way

they can, taking the least risk. If an attacker knows little about a system, he cannot give an investor good estimates about the cost of a successful attack, discouraging investment, diverting it towards scams that are better understood.

Obscurity can increase the overall cost of an attack, as the attacker spends time gathering information about the system. It can delay the date of ultimate success of an attack, decreasing the payback period in systems that are periodically renewed.

Risk of using obscurity is low if the system relies on well-known algorithms and techniques, while implementation details are obscured and kept proprietary. Thus, a system designer can take advantage of widespread analysis of various crucial aspects of a system without necessarily revealing much about it. There is residual risk from this compromise. Using any obscurity bars the designer from unsolicited, but helpful observations by others. It is therefore important that an extensive and continuous private review and audit process compensate this policy. Obscurity should never be viewed as an explicit countermeasure. Finally, obscurity should be used as a means of increasing leverage over the intruder, not as the fundamental countermeasure.

## 9. Final Perspective

New interest by researchers in security fault analysis and hostile reverse engineering is desired and helpful. It is necessary for the design community and the research community to better understand one another, and to share process and vocabulary. Attacks on system implementations should not be characterized as generic attacks on algorithms (RSA and DES have not been broken). Research can help designers learn of new threats and designers can help researchers understand the practicality of those attacks and how threats can be countered and risks reduced. It is necessary for all parties to be more careful in describing their abilities and results. In the area of security, advertising hype by product organizations and poorly bounded claims by researchers are both counterproductive. Great amounts of time have been wasted because busy people have to explain to executives that no, RSA and triple DES have not been broken. At best, we can say that the new attacks break hypothetical, simplified implementations of systems that use RSA, DES, and similar algorithms. It's better to spend time analyzing the legitimate attacks in the proper context than correcting mis-impressions caused by hype and inadequate vocabulary.

Security by obscurity strategies need to be tuned so that feedback by the research and design communities can be most useful while preserving the benefit of increased risk and work factor. Professional secure systems designers always assume that their systems will be broken and that any given countermeasure will be rendered ineffective. However, this does not imply that a given system cannot be designed to do its job and to limit and manage the overall security risk.

Finally, I do not mean to argue that the systems defended here are indeed secure. Reams of paper would be required to do that. These systems are bound to be

successfully attacked if they are commercially successful. However, I believe that the success of attackers can be limited.

## 10. References


[AK96a] RJ Anderson and M Kuhn, "Tamper Resistance – a Cautionary Note." *Proceedings of the 2nd Workshop on Electronic Commerce*, Oakland, CA, November 18-20, 1996.

[AK96b] RJ Anderson and M Kuhn, "Warning to the Crypto and Banking Committee – A serious weakness of DES", Draft – November 2, 1996. Ftp://ftp.cl.cam.ac.uk/users/rja14/warning

[AK96c] RJ Anderson and M Kuhn, "Improved Differential Fault Analysis", Draft, November 1996 from Ftp://ftp.cl.cam.ac.uk/users/rja14/dfa

[BDL96] D Boneh, RA DeMillo, RJ Lipton "Cryptanalysis in the presence of Hardware Faults'" Preprint – Sept, 1996. Re-issued as "On the importance of Checking Computations", preprint – 1996, to appear at Eurocrypt '97.

[Bel96a] Bellcore Security Alert, "Now Smart Cards can leak Secrets – A new breed of Crypto Attack on 'Tamperproof' Tokens Cracks Even the Strongest RSA Code", September 1995.

[Bel96b] Bellcore Press Release, http://www.bellcore.com/PRESS/ADVSRY96/ smrtcrd.html , Sept. 1996

[BS96a] E Biham, A Shamir, "A new cryptanalytic attack on DES", preprint Oct. 18, 1996

[BS96b] E Biham, A Shamir, "Differential Fault Analysis: Identifying the Structure of Unknown Ciphers Sealed in Tamper-Proof Devices", Preprint November 11, 1996.

[CFPR96] D Coppersmith, M Franklin, J Patarin, M Reiter, "Low Exponent RSA with related messages" Eurocrypt '96.

[FIPS] Federal Information Processing Standard 140-1, National Institute of Standards and Technology

[HM95] DN Heer and DP Maher, "The Heart of the New Information Appliance", IEEE Transactions on Consumer Electronics, August, 1995.

[LEN96] Arjen Lenstra, Citibank internal memo

[LT96] London Telegraph, November 19, 1996, Page 1

[Mc94] John McCormac, *European Scrambling Systems - Circuits, Tactics, and Techniques*, Waterford University Press, 1994

[NYT96] New York Times, September 25, 1996, Page 1, business section.