Abstract:

Today’s Smartcard- and Security-IC’s are no longer just phone cards, they are now embedded cryptographic security processors and have to face special challenges concerning architecture, design methodology and technology. The requirements for power consumption and performance are quite contradictory. On the one hand, embedded Java capability needs the high computing power of a 16 bit or even 32 bit CPU core. On the other hand the contactless mode of operation requires lowest power ability. Smartcard IC’s have specific co-processors for efficient execution of several cryptographic algorithms and a set of peripherals to enable a flexible use of the controller for many kinds of applications. The most important challenge of security IC’s is their resistance against attacks. An overview of invasive and non-invasive attacks and examples for appropriate countermeasures in system and circuit design will be given.

1. Introduction

Security is a basic feature of smartcards that requires dedicated concepts with impacts on circuit and system design. There are various attacks the system must be resistant against. Section 2 reviews attacks and gives brief descriptions.

Section 3 describes the working principles and countermeasures for some selected attacks in more detail. Typically there is a tradeoff between the security mechanisms on the one hand and the high performance and low power requirements on the other hand.

Section 4 will close the paper with a short introduction to certification and a discussion of the presented aspects.

2. Overview on Attacks

There is a natural differentiation in two classes of attacks: invasive and non-invasive attacks [21]. An attack that does not destroy the circuit or its working capability is called a non-invasive attack.

Examples are software attacks that use the regular communication interface of the processor and exploit security vulnerabilities of protocols, cryptographic algorithms, or implementation.

Side-channel attacks (SCA), on the other hand, exploit additional information leaked during the operation of the system. As depicted in Fig. 1 the attacker will use the input \( m \) (called message text) and the output \( c \) (called cipher text) as well as the eavesdropped side channel information \( SCI \) to retrieve the secret information \( K \) (denoted as key).

Side-channel information can be contained in the characteristic power consumption, the timing, or the electromagnetic emanation of the device during the processing of secret information.

Timing attacks (TA) have been introduced in [19] to break implementations of RSA and several digital signature schemes. TA exploits dependencies of the runtime of an algorithm on the processed secret information.

Power analysis attacks, on the other hand, exploit the fact that, in general, the instantaneous power consumption of the circuit depends on the data being processed by the circuit. The effect is prominent especially in the widely used CMOS design style. Differential Power Analysis (DPA), first introduced in [20], allows the attacker to exploit correlations between the observable instantaneous power consumption and intermediate results involving the secret. While in a Simple Power Analysis (SPA) the characteristic of a single power trace is evaluated, DPA applies a statistical analysis on a collection of power traces for several chosen input values. During the last years it became more and more obvious that it is extremely difficult to protect a security device against DPA [1], [2], [6], [8], [9], [11], [14], [18], [23], [24], [31], [32], [33], [34].

In the spirit of power analysis attacks Electromagnetic Emanation Analysis (EMA) extracts secret information from the electro-magnetic radiation emitted during the operation of the device [13]. In analogy to the
power analysis SEMA is based on a single trace of the field magnitude, whereas DEMA denotes the differential analysis based on a set of traces. In contrast to SPA and DPA the EMA yields also a spatial resolution of the leakage signal [12]. As a consequence the signal-to-noise ratio of the side-channel signal can be significantly increased and also the application of higher order differential attacks can be facilitated.

Other kind of non-invasive attacks are the so-called **Fault Attacks** (FA). In these attacks temporary faults are induced during the processing of secret information. In [7] it was shown that transient faults induced during the RSA computation can be used to retrieve the secret key by analyzing the faulty result. Also symmetric algorithms, like the DES [4] and the AES [27] encryption standard, have been attacked successfully by toggling few bits in an intermediate round of an encryption and applying some cryptanalysis to the observed faulty results. Fault induction can even lead to seemingly trivial attacks on a device: faults in program counters, loop counters, or branch conditions can lead to extended runtime of loops, e.g., forcing normal port I/O routines to output data or program code outside the regular output buffer memory. Reduced loop counts can convert a secure iterated block cipher into a single-round variant, which can be easily broken to obtain the secret key [3]. A recent overview on FA and proposals for countermeasures can be found in [5]. In the next section we describe various fault induction mechanisms.

One of the first reviews on hardware techniques for breaking into smartcards [21] describes **invasive attacks**, such as methods for depackaging, layout reconstruction, micro-probing, or particle beam techniques to modify the chip stack.

Accessing a single wire of a secure device by probing it with a needle and observing the data transfer can already break the cryptographic system. Unprotected implementations of secret key encryption schemes, like DES, as well as public key schemes, like RSA, are susceptible to this type of attack [15]. A mathematical analysis and a construction principle for circuits resistant against probing are given in [16].

### 3. Attacks and Countermeasures

#### 3.1 Power-Analysis

**3.1.1 Ad-hoc Approaches**

The first class of **ad-hoc approaches** against power-analysis attacks tries to reduce the signal-to-noise ratio of the side-channel leakage and finally to hide the usable information in the noise. Suggested methods are detached power supplies [29], the addition of power noise generators, or the application of a probabilistic disarrangement of the times at which the attacked intermediate results are processed. The latter can be achieved by inserting random delays or applying randomizations to the execution path [21]. While such measures certainly increase the experimental and computational working load of the attacker they do not render the attack infeasible. In practice, typically several countermeasures are combined [9]. This can reduce the correlations down to a level that makes a DPA practically impossible. However, higher order differential attacks or the possibility of obtaining a spatial resolution of the power consumption by observing local electromagnetic emissions may again open a backdoor for professional attackers.

#### 3.1.2 Circuit Design Approaches

The second class of countermeasures aims at removing the root cause for side-channel leakage information. In standard CMOS style circuits, for instance, the power consumption depends strongly on the processed data [28]. An individual gate consumes power from the supply on a 0-1 output transition. During the 1-0 output transition the stored energy is dissipated, and in the degenerate cases of 0-0 or 1-1 transitions no energy is used (leakage currents are neglected here). The goal of special circuit design styles is to make the power consumption of individual logic gates independent of the values of the input signals or the Hamming distance between subsequent input signals.

The **Sense Amplifier Based Logic (SABL)**, e.g., is a dynamic and differential logic [31] [32] which has one switching event per cycle independent of the input value, i.e., also in the degenerate case in which the gate does not change the logical state. Generally, in a differential logic a signal is encoded on more than one wire. In dynamic logic evaluation and precharge phases alternate. Hence, in a dual-rail dynamic differential logic, like DCVSL (differential cascode voltage switching logic) [28] or SABL, exactly one output wire is 0 in the evaluation phase and both output wires are charged to the same value (0 or 1) in the precharge phase. An example for a DCVSL XOR gate is shown in Fig. 3. In fact, all logic styles currently proposed to counteract DPA are based on these principles. Data independent (constant) power consumption requires a maximum activity factor, and hence maximum power consumption. SABL (or other dual-rail differential logic) circuits leak less side-channel information than CMOS circuits, i.e. for a successful DPA more power traces must be evaluated.

---

**Figure 2: Taxonomy of attacks.**
To achieve a constant power consumption it is essential that the load capacitances of the differential outputs are matched, i.e. the intrinsic gate output capacitances and the interconnect capacitances must match. However, remaining asymmetries (e.g. parasitic, cross-coupling) make a DPA still possible. Disadvantages of this circuit style are the lack of standard cell libraries and automated routing tools for matching the interconnect capacitances. This leads to a full-custom design style. Area (power consumption) of a SABL circuit design are approximately 3.5 times (4.5 times) larger than for a corresponding CMOS design. The performance is reduced by a factor of two due to the two-cycle clocking scheme.

Figure 3: Example for a DCVSL XOR-gate.

The wave dynamic differential logic style (WDDL) adopts the ideas of SABL. It implements the behavior of a dynamic and differential logic, but is based on standard CMOS cells [33]. Area and power consumption are approximately 3.5 times larger than for a CMOS design. The performance is two times smaller.

Self-timed asynchronous circuits seem to have appealing advantages for secure systems at the first glance. Experimental results [12] indicate that the information leakage in DPA attacks is smaller than in a corresponding synchronous circuit. However, the reduction is not sufficient to protect against DPA. In EMA measurements the absence of the clock even facilitates the attack, because noisy components due to the clock signal are absent. The commonly used dual-rail encoding (or generally a 1-of-"n" encoding) does not guarantee that the consumed power is independent of the value of the data. The reason is, like in the case of SABL, that any imbalance in the gate and interconnect load capacitances, due to asymmetries in the routing will lead to leakage of information. Additionally, variations in interconnect and gate capacitances can also lead to leaking of information in the time domain. The advantage of asynchronous circuits should be the low power consumption compared to CMOS circuits. Furthermore, the redundant dual-rail data encoding offers the possibility to add checker circuits for the forbidden code word, thus increasing the robustness in fault attacks.

3.1.3 Masking Approaches

The third class of measures counteracts DPA by randomizing intermediate results occurring during the execution of the cryptographic algorithm. The idea behind this approach is that the power consumption of operations on randomized intermediate data should not be correlated with the actual plain intermediate data. Algorithmic countermeasures in the context of symmetric ciphers based on secret sharing schemes have been independently proposed in [8] and [14].

Masking at algorithm level for asymmetric algorithms [25], as well as the symmetric algorithms, DES and AES, have been developed [2] [6]. Cryptographic algorithms often combine Boolean functions (like logical XOR or AND operations) and arithmetic functions (operations in fields with characteristic bigger than two). Masking operations for these two types of functions are referred to as Boolean and arithmetic masking, respectively. This poses the problem of a secure conversion between the two types of masking in both directions [2].

It is appealing to apply the idea of randomizing intermediate results already on the level of logic gates. Masking at gate level leads to circuits where no wire carries a value which is correlated to an intermediate result of the algorithm. Clearly this approach is more generic than the algorithmic approach. Masking at gate level is independent of the specifically implemented algorithm. Once a secure masking scheme has been developed the generation of the masked circuit from the algorithm can be automated and a computer program can convert the digital circuit of any cryptographic algorithm to a circuit of masked gates. This would also relieve the designers or implementers of cryptographic algorithms from the complex task of elaborating a specific solution against side-channel leakage for each new algorithm or new implementation variant of the algorithm. Various generic masking schemes have been proposed. In [24] the multiplexor gate (MUX) used in the implementation of nonlinear operations, like S-boxes, is replaced by a masked MUX gate which in turn consists of three MUX gates. In [18] the basic operations of an arithmetic-logic unit are protected with one or more random masks attached to each masked gate. In [34] correction terms for the AND gate in the nonlinear components of the S-Boxes of the AES are introduced. It has been shown that it is possible to break masking schemes that rely on one mask using advanced DPA methods [1].

The random switching logic (RLS) proposed in [30] uses a random input per gate and introduces an enable signal which forces the output to a definite value until all input signals are stable. Hence it is also a hidden two-cycle scheme, however, requiring a delicate adjustment of the timing of the enable signals.

The security analyses of masking schemes, conducted so far, were based on the implicit assumption that the input signals of any (masked) gate in a combinational CMOS circuit arrive at the same time. Recently it has been shown [23], that this assumption is not true: the output of the gate possibly switches several times during one clock cycle. The transitions at the
output of a gate, previous to the stable state right before the next clock edge is attained, are known as glitches. Glitches are a typical phenomenon in CMOS circuits and extensively discussed in the literature on VLSI design [28]. Because a glitch can cause a full swing transition at the output of the gate, just like the ‘proper’ transition to the final value, a glitch is not a negligible higher order effect. As made evident in [23] glitches do not just add a background noise due to uncorrelated switching activity – the dissipated energy of nonlinear masked gates is correlated to the processed values whenever the input values do not arrive simultaneously (forcing the output of the gate to toggle several times). Hence glitches can carry side-channel information and their effect must be included in the analysis of any secure masking scheme [11].

3.2 Fault Analysis

The most common techniques for injecting faults into a system are:

- **Spike attacks**
  The power supply and the clock as well as all input/output signals of the system can be used for spike or glitch penetration.

- **Light attacks**
  Light within a wide range of frequencies and intensities can be applied globally (on the whole chip) or locally (only on a small area). Flash-lights or lasers are commonly used. The induced photoelectric current can lead to faulty switching events.

- **Ionizing radiation attacks**
  Alpha particles, ionizing ion beams or X-rays can be applied to generate single or multiple event upsets (SEU, MEU).

- **Temperature, voltage, or frequency variation**
  Generally the attacker will try to operate the chip out of the specified operation range to trigger faulty behavior.

For further reading we refer to [5] and [21]. It is especially challenging to detect or prevent faults in a secure system with restricted area and power resources. We can distinguish between active protection and passive protection measures. Measures of the first category will try to recognize the penetration or prevent faults in advance, while those in the second category will try to detect the effect of faults and react post failure. Examples for penetration detection are sensors which monitor the operating conditions, such as temperature, clock frequency, and voltage, or filters which remove voltage spikes applied to the external contacts.

The class of passive protection measures comprises various redundancy schemes. Hardware redundancy schemes are commonly applied to memories. Error detection or correction codes are capable to detect or correct errors in non-volatile memories (e.g. EEPROM, Flash), and internal or external RAM’s. The protection of data paths in CPU’s usually requires a larger overhead in terms of area and power. We mention duplication schemes, e.g. triple modular redundancy (TMR), redundant residue number systems (RRNS), parity or Berger codes, or modular redundancy codes (e.g. modulo-3 checkers) [22]. For cryptographic algorithms there are several dedicated methods which are more efficient than the generic protection schemes. Time redundancy schemes, on the other hand, avoid the hardware overhead by repeating the calculation or parts of it once or several times. Avoiding the identical data path by a transformation of the input value in the repeated calculation can significantly increase the working load of the attacker or render the attack impossible.

3.3 Probing

Probing can be done with needles that are placed on wires of the circuit. With the decreasing minimum feature size of modern manufacturing technologies probing becomes increasingly difficult, especially if more than one probing needle is used. But, nevertheless, this type of attack is still a serious threat especially if focused ion beam technique (FIB) is used to connect test pads. Additionally there are other methods not requiring a direct contact to get information about circuit activity. Accessing and observing the data transfer on a single wire of a secure device can already break the cryptographic system [15]. For a theory of securing a circuit at the gate level against attacks, focused on probing, we refer to [16]. Probing needles can also be used to induce faults by applying an electric potential to the needle. A passive countermeasure against probing is the on-the-fly encryption of memories and of information channels carrying secret information (like system buses). Active countermeasures against probing are shield structures which cover areas containing secret information.

3.4 Reverse Engineering

Attacks of this class try to find out information of the system by destroying it. If an attacker can recognize the internal structure of the circuit (e.g. the connections in a ROM memory), he can access secret information (like keys stored in the ROM). The chip is reengineered and the different layers of the technology are taken off and analyzed. The result is a net list of the circuit or of parts of the chip. A powerful countermeasure is the encryption of the memories using an established encryption algorithm with an appropriate key length.

4. Discussion

As it is indicated in the preceding chapters secure systems require high effort in terms of concept and implementation tasks as well as chip area and power consumption. Chip manufacturers have to spend a large amount of money for development and production. Therefore it is important that the security level of such a system is independently investigated and reliably documented. This is done by certification bodies. The best known official certificates are those from ITSEC (Information Technology Security Evaluation
Criteria) and CC (Common Criteria). Additionally there are so-called “type approvals”, which are defined by banking card issuers like Mondex, Proton, EMV (Europay, MasterCard and Visa), as well as American Express and ZKA (Zentraler Kreditausschuss). A proven security system has to get a certificate or pass a type approval if it should be applied in specific applications. The certification is based on a rating of resistance against attacks. On the one hand the evaluation depth is stated in a number (EALx). On the other hand the security level is given. For CC this level can be “basic”, “medium” or “high” and depends on the total number of points a device gets during investigation.

The certification is based on a rating of resistance against attacks. On the one hand the evaluation depth is stated in a number (EALx). On the other hand the security level is given. For CC this level can be “basic”, “medium” or “high” and depends on the total number of points a device gets during investigation.

<table>
<thead>
<tr>
<th>Common Criteria</th>
<th>Limit for Hardware</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAL 1</td>
<td>Functionally tested</td>
</tr>
<tr>
<td>EAL 2</td>
<td>Semiformally verified and checked</td>
</tr>
<tr>
<td>EAL 3</td>
<td>Methodically designed, tested and reviewed</td>
</tr>
<tr>
<td>EAL 4</td>
<td>Methodically tested, and checked</td>
</tr>
<tr>
<td>EAL 5</td>
<td>Structurally tested</td>
</tr>
<tr>
<td>EAL 6</td>
<td>Semiformally design and tested</td>
</tr>
<tr>
<td>EAL 7</td>
<td>Formally verified design and tested</td>
</tr>
</tbody>
</table>

The resulting security level is stated following the evaluation depth.

Figure 4: Common Criteria EAL levels

There are five criteria for the rating which are the required levels of expertise, the required level of knowledge about the device, the required number of devices needed for the attack, the equipment needed, and the amount of time needed for the attack. For these criteria the ratings for the identification and the exploitation of the attack are accumulated. The certification is a long lasting and expensive process, but at the end the customer buys a product, that has a proven security level required for his application. Details of the certification procedure are given e.g. in [10] and [17].

The kind of attack that is investigated and rated during the process is subject to change. Indeed the certification body will go for that attack that has the highest potential of successfully breaking the system. This means that the security system developer always has to keep in mind that there might be more sophisticated attacks in the future.

Ab initio theoretically provable security does not seem to be possible and the only chance a security system designer has to build a successful and long lasting product is to think about potential attacks long before the system is designed and implemented.

References:

[14] L. Goubin and J. Patarin: DES and Differential Power Analysis — The Duplication Method, Crypto-
ographic Hardware and Embedded Systems — CHES 1999, (Ç. K. Koç, C. Paar, eds.), Lecture Notes in
Computer Science, vol. 1717, pp. 158-172, Springer-
Verlag, 1999.

Attacks on Tamper-Resistant Devices, Crypto-
graphic Hardware and Embedded Systems — CHES 1999,
(Ç. K. Koç, C. Paar, eds.), Lecture Notes in Compu-
ter Science, vol. 1717, pp. 303-315, Springer-
Verlag, 1999.

[16] Y. Ishai, A. Sahai, and D. Wagner: Private Circuits:
H. Handschuh, P. Paillier, and J. Stern: Probing
Attacks on Tamper-Resistant Devices, Crypto-
graphic Hardware and Embedded Systems — CHES 1999,
(Ç. K. Koç, C. Paar, eds.), Lecture Notes in Compu-
ter Science, vol. 1717, pp. 303-315, Springer-
Verlag, 1999.

Potential to Smartcards, Version 1.1, July 2002,
available at
http://www.commoncriteriaportal.org/

[18] F. Klug, O. Kniffel, B. Gammel: Rechenwerk, Ver-
fahren zum Ausführen einer Operation mit verschlüs-
selten Operanden, Carry-Select-Addierer und Krypto-
graphieprozessor, German Patent DE 10201449 C1,

Advances in Cryptology — CRYPTO 1996, (N. Kobilitz, ed.), Lecture Notes in Computer Science,

Power Analysis, Advances in Cryptology — CRYPTO '99,
(M. J. Wiener, ed.), Lecture Notes in Computer Science,

[21] O. Kömerling and M. G. Kuhn: Design principles for Tamper-Resistant Smartcard Processors, Proceed-
ings of the USENIX Workshop on Smartcard Tech-


[23] S. Mangard, T. Popp, and B. M. Gammel: Side-
Channel Leakage of Masked CMOS Gates, Topics in

Method and Apparatus for Preventing Information
Leakage Attacks on a Microelectronic Assembly, US

Power Analysis Attacks of Modular Exponentiation
in Smartcards, Cryptographic Hardware and Embed-
ded Systems — CHES 1999, (Ç. K. Koç, C. Paar,
ed.), Lecture Notes in Computer Science, vol. 1717,

[26] T. S. Messerges, E. A. Dabbish, and R. H. Sloan:
Examining Smart-Card Security under the Threat of
Power Analysis Attacks, IEEE Transactions on Com-

[27] G. Piret and J. J. Quisquater: A Differential Fault
Attack Technique Against SPN Structure, with
Application to the AES and KHAZAD, Crypto-
graphic Hardware and Embedded Systems — CHES
2003 (C. D. Walter, Ç. K. Koç, C. Paar, eds.), Lecture
Notes in Computer Science, vol. 2779, pp. 77-88,


[29] A. Shamir: Protecting Smart Cards from Passive
Power Analysis with Detached Power Supplies,
Cryptographic Hardware and Embedded Systems —
CHES 2000, (Ç. K. Koç, C. Paar, eds.), Lecture
Notes in Computer Science, vol. 1965, pp. 71-77,

Switching Logic: A Countermeasure against DPA
based on Transition Probability, Cryptology ePrint
eprint.iacr.org).

[31] K. Tiri, M. Akmal, and I. Verbauwhede: A Dyna-
mic and Differential CMOS Logic with Signal Inde-
dependent Power Consumption to Withstand Differen-
tial Power Analysis on Smart Cards, Proceedings of
28th European Solid-State Circuits Conference —

Algorithms against DPA at the Logic Level: Next
Generation Smart Card Technology, Cryptographic
Hardware and Embedded Systems — CHES 2003,
(C. D. Walter, Ç. K. Koç, C. Paar, eds.), Lecture
Notes in Computer Science, vol. 2779, pp. 137-151,

[33] K. Tiri and I. Verbauwhede: A Logic Level Design
Methodology for a Secure DPA Resistant ASIC or
FPGA Implementation, Proceedings of Design,
Automation and Test in Europe Conference — DATE

[34] E. Trichina, D. De Setta, and L. Germani: Simplified
Adaptive Multiplicative Masking for AES, Crypto-
graphic Hardware and Embedded Systems — CHES
2002, (B. S. Kaliski Jr., Ç. K. Koç, C. Paar, eds.),