Tamper resistance and hardware security

Dr Sergei Skorobogatov

http://www.cl.cam.ac.uk/~sps32       email: sps32@cam.ac.uk

UNIVERSITY OF CAMBRIDGE
Computer Laboratory
Talk Outline

• Introduction
• Attack awareness
• Tamper protection levels
• Attack methods
  – Non-invasive
  – Invasive
  – Semi-invasive
• Protection against attacks
• Conclusions
• Slides
  – http://www.cl.cam.ac.uk/~sps32/PartII_030214.pdf
Introduction

• Protection of systems and devices against physical attacks
  – protecting secrets from being stolen
  – preventing unauthorised access
  – protecting intellectual property from piracy
  – preventing fraud

• Examples
  – locks and sensors to prevent physical access
  – smartcards to hold valuable data and secret keys
  – electronic keys, access cards and hardware dongles
  – electronic meters, SIM cards, PayTV smartcards
  – crypto-processors and crypto-modules for encryption
  – mobile phones, game consoles and many other devices
Why do we need hardware security?

• Theft of service
  – attacks on service providers (satellite TV, electronic meters, access cards, software protection dongles)

• Access to information
  – information recovery and extraction
  – gaining trade secrets (IP piracy)
  – ID theft

• Cloning and overbuilding
  – copying for making profit without investment in development
  – low-cost mass production by subcontractors

• Denial of service
  – dishonest competition
  – electronic warfare
Who need secure chips?

• There is growing demand for secure chips
  – car industry, service providers, manufacturers of various devices
  – banking industry and military applications

• Technical progress pushed secure semiconductor chips towards ubiquity
  – consumer electronics (authentication, copy protection)
  – aftermarket control (spare parts, accessories)
  – access control (RF tags, cards, tokens and protection dongles)
  – service control (mobile phones, satellite TV, license dongles)
  – intellectual property (IP) protection (software, algorithms, design)

• Challenges
  – How to design secure system? (hardware security engineering)
  – How to evaluate protection? (estimate cost of breaking)
  – How to find the best solution? (minimum time and money)
How to design a secure system?

• What is the reason to attack your system?
  – attack scenarios and motivations: theft, access, cloning or DoS

• Who is likely to attacks your system?
  – classes of attackers: outsiders, insiders or funded organisations

• What tools would they use for the attacks?
  – attack categories: side-channel, fault, probing, reverse engineering
  – attack methods: non-invasive, invasive, semi-invasive

• How to protect against these attacks?
  – estimate the threat: understand motivation, cost and probability
  – develop adequate protection by locating weak points
  – perform security evaluation
  – choose secure components for your system (blocks and chips)
Choosing secure components

• What has changed in the past?
  – too many designs and devices on the market
  – vast majority of devices are claimed to be secure
  – security started to be used for marketing purposes
  – virtually impossible to test everything

• What are the problems?
  – certification does not provide guarantee against attacks
  – manufacturers do not carry any obligations or legal responsibility
  – no such thing as security benchmark
  – no ways of comparing devices from different manufacturers
  – no chip manufacturer will tell you the truth about security

• Need for security educated system engineers
Attack categories

- Side-channel attacks
  - techniques that allows the attacker to monitor the analog characteristics of supply and interface connections and any electromagnetic radiation

- Software attacks
  - use the normal communication interface and exploit security vulnerabilities found in the protocols, cryptographic algorithms, or their implementation

- Fault generation
  - use abnormal environmental conditions to generate malfunctions in the system that provide additional access

- Microprobing
  - can be used to access the chip surface directly, so we can observe, manipulate, and interfere with the device

- Reverse engineering
  - used to understand the inner structure of the device and learn or emulate its functionality; requires the use of the same technology available to semiconductor manufacturers and gives similar capabilities to the attacker
Attack methods

• Non-invasive attacks (low-cost)
  – observe or manipulate with the device without physical harm to it
  – require only moderately sophisticated equipment and knowledge to implement

• Invasive attacks (expensive)
  – almost unlimited capabilities to extract information from chips and understand their functionality
  – normally require expensive equipment, knowledgeable attackers and time

• Semi-invasive attacks (affordable)
  – semiconductor chip is depackaged but the internal structure of it remains intact
  – fill the gap between non-invasive and invasive types, being both inexpensive and easily repeatable
Tamper protection levels

• Level ZERO (no special protection)
  – microcontroller or FPGA with external ROM
  – no special security features are used. All parts have free access and can be easily investigated
  – very low cost, attack time: minutes to hours

D.G. Abraham et al. (IBM), 1991
Tamper protection levels

• Level LOW
  – microcontrollers with proprietary access algorithm, remarked ICs
  – some security features are used but they can be relatively easy defeated with minimum tools required
  – low cost, attack time: hours to days
Tamper protection levels

• Level MODL
  – microcontrollers with security protection, low-cost hardware dongles
  – protection against many low-cost attacks; relatively inexpensive tools are required for attack, but some knowledge is necessary
  – moderate cost, attack time: days to weeks
Tamper protection levels

• Level MOD
  – smartcards, high-security microcontrollers, ASICs, CPLDs, hardware dongles, i-Buttons, secure memory chips
  – special tools and equipment are required for successful attack as well as some special skills and knowledge
  – high cost, attack time: weeks to months
Tamper protection levels

- **Level MODH**
  - secure i-Buttons, secure FPGAs, high-end smartcards, ASICs, custom secure ICs
  - special attention is paid to design of the security protection; equipment is available but is expensive to buy and operate
  - very high cost, attack time: months to years

![Tamper protection components](image)
Tamper protection levels

• Level HIGH
  - military and bank equipment
  - all known attacks are defeated. Some research by a team of specialists is necessary to find a new attack
  - extremely high cost, attack time: years
Tamper protection levels

• Division into levels from ZERO to HIGH is relative
  – some products designed to be very secure might have flaws
  – some products not designed to be secure might still end up being very difficult to attack
  – technological progress opens doors to less expensive attacks, thus reducing the protection level of some products

• Proper security evaluation must be carried out to estimate whether products comply with all the requirements
  – design overview for any possible security flaws
  – test products against known attacks
Non-invasive attacks

• Non-penetrative to the attacked device
  – normally do not leave tamper evidence of the attack

• Tools
  – digital multimeter
  – IC soldering/desoldering station
  – universal programmer and IC tester
  – oscilloscope, logic analyser, signal generator
  – programmable power supplies
  – PC with data acquisition board, FPGA board, prototyping boards

• Types of non-invasive attacks: passive and active
  – side-channel attacks: timing, power and emission analysis
  – data remanence
  – fault injection: glitching, bumping
  – brute forcing
Non-invasive attacks: side-channel

- Timing attacks aimed at different computation time
  - incorrect password verification: termination on incorrect byte, different computation length for incorrect bytes
  - incorrect implementation of encryption algorithms: performance optimisation, cache memory usage, non-fixed time operations

- Power analysis: measuring power consumption in time
  - very simple set of equipment – a PC with an oscilloscope and a small resistor in power supply line; very effective against many cryptographic algorithms and password verification schemes
  - some knowledge in electrical engineering and digital signal processing is required
  - two basic methods: simple (SPA) and differential (DPA)

- Electro-magnetic analysis (EMA): measuring emission
  - similar to power analysis, but instead of resistor, a small magnetic coil is used allowing precise positioning over the chip
Non-invasive attacks: power analysis

- Simple power analysis (SPA): difference in instruction flow
  - 8-byte password check in Freescale MC908AZ60A microcontroller
  - 1 byte at a time, 1 of 256 attempts leads to distinctive power trace
  - full password recovery in 2048 attempts (less than 10 minutes)

```
loop:     CBEQX #$FE, ptr3 ; check for end
         JSR sub_recv ; receive byte
         CBEQ X+, ptr2 ; compare byte
         CLR adr_50 ; clear status

ptr1:    BRA loop ; loop
ptr2:    BRA ptr1 ; time alignment
ptr3:    LDX #$FF ; set address
         LDA adr_50 ; check status
         BEQ cont ; skip flash enable
         STX , X ; flash enable

cont:     … … …
```
Non-invasive attacks: power analysis

- Differential power analysis (DPA): correlation with secret
  - AES decryption in asynchronous ASIC (130 nm, 1.5V), 128-bit key
  - first round of decryption starts with XORing the input data with round key, the difference is only in the input data and the result
  - full key recovery in 256 attempts with each attempt requiring average of 4096 traces (~2 minutes per attempt, total 8 hours)
Non-invasive attacks: data remanence

• Data remanence in SRAM
  – residual representation of data after erasure – first discovered in magnetic media then appeared to be the case for other memories
  – low temperature data remanence is dangerous to tamper resistant devices which store keys and secret data in a battery backed-up SRAM
  – long period of time data storage causes the data to be “burned-in” and likely to appear after power up; dangerous to secure devices which store keys at the same memory location for years

• Eight SRAM samples were tested at different conditions
  – at room temperature the retention time varies from 0.1 to 10 sec
  – cooling down to −20°C increases the retention time to 1…1000 sec, while at −50°C the data retention time is 10 sec to 10 hours
  – grounding the power supply pin reduces the retention time
Non-invasive attacks: data remanence

• Data remanence in non-volatile memories

• EPROM, EEPROM and Flash
  – widely used in microcontrollers and smartcards
  – use floating-gate transistors for storage, $10^3 - 10^5$ e$^-$

• Levels of remanence threat
  – file system (erasing a file – undelete)
  – file backup (software features)
  – smart memory (hardware buffers)
  – memory cell

• Possible outcomes
  – circumvention of security in microcontrollers, FPGAs, smartcards
  – information leakage through shared EEPROM and Flash areas between different applications in secure chips
Non-invasive attacks: data remanence

- Data remanence in EEPROM and Flash
  - threshold voltage of a memory cell ($V_{TH}$) is compared with reference voltage which is proportional to the power supply and can be influenced
  - memory bulk erase cycles
    - Flash memory, after 100 erase cycles: $\Delta V_{TH} = 100 \text{ mV}$
    - EEPROM memory, after 10 erase cycles: $\Delta V_{TH} = 1 \text{ mV}$
  - information successfully recovered from PIC16F84 after 10 erase cycles

![Threshold Voltage Change During Erase Cycles](image-url)
Non-invasive attacks: fault injection

- Glitch attacks
  - clock glitches
  - power supply glitches
  - data corruption

- Security fuse verification in the Mask ROM bootloader of the Motorola MC68HC05B6 microcontroller
  - double frequency clock glitch causes incorrect instruction fetch
  - low-voltage power glitch results in corrupted EEPROM data read

```
LDA #01h
AND $0100 ; the contents of the EEPROM byte is checked
loop:
BEQ loop ; endless loop if bit 0 is zero
BRCLR 4, $0003, cont ; test mode of operation
JMP $0000 ; direct jump to the preset address
cont: … … …
```
Non-invasive attacks: fault injection

- Data protection with integrity check
  - verify memory integrity without compromising confidentiality
  - How secure is the “No Readback” solution?
Non-invasive attacks: fault injection

• Authentication using encryption
  – verify if a user knows the secret key by asking him to encrypt a message with his key
  – How secure is the 'No Readback' scheme against key extraction?

Readback access controlled by security fuse

No Readback access only secure verification
Non-invasive attacks: fault injection

- Bumping and selective bumping attacks
  - aimed at internal integrity check procedure on a chip (verification and authentication using encryption or hash functions)
  - aimed at blocks of data down to bus width or at individual bits within the bus

- Power supply glitching attack on secure microcontroller
  - exhaustive search: $2^{127}$ attempts per 128-bit AES key $\rightarrow$ >trillion years
  - bumping: $2^{15}$ attempts per 16-bit word, 100ms cycle, 8 hours for AES key
  - selective bumping: $2^7$ attempts per 16-bit word, 2 minutes for AES key
Non-invasive attacks: brute forcing

• Brute force attacks
  – searching for keys and passwords, exploiting inefficient selection of keys and passwords
  – recovering design from CPLDs, FPGAs and ASICs
  – eavesdropping on communication to find hidden functions
  – applying random signals and commands to find hidden functionality

• Modern chips deter most brute force attacks
  – longer keys make searching infeasible
  – moving from 8-bit base to 32-bit base means longer search
  – CPLDs, FPGAs and ASICs became too complex to analyse
  – too large search field for finding hidden functionality
Invasive attacks

• Penetrative attacks
  – leave tamper evidence of the attack or even destroy the device

• Tools
  – IC soldering/desoldering station
  – simple chemical lab
  – high-resolution optical microscope
  – wire bonding machine, laser cutting system, microprobing station
  – oscilloscope, logic analyser, signal generator
  – scanning electron microscope and focused ion beam workstation

• Types of invasive attacks: passive and active
  – decapsulation, optical imaging, reverse engineering
  – microprobing and internal fault injection
  – chip modification
Invasive attacks: sample preparation

- Decapsulation
  - manual with fuming nitric acid (HNO$_3$) and acetone at 60°C
  - automatic using mixture of HNO$_3$ and H$_2$SO$_4$
  - full or partial
  - from front side and from rear side

- Challenging process for small and BGA packages
Invasive attacks: imaging

• Optical imaging
  – resolution is limited by optics and wavelength of a light:
    \[ R = 0.61 \frac{\lambda}{NA} = 0.61 \frac{\lambda}{n \sin(\mu)} \]
    • reduce wavelength of the light using UV sources
    • increasing the angular aperture, e.g. dry objectives have \( NA = 0.95 \)
    • increase refraction index of the media using immersion oil (\( n = 1.5 \))
Invasive attacks: imaging

• Optical imaging
  – image quality depends on microscope optics
    • depth of focus helps in separating the layers
    • geometric distortions pose problem for later post-processing
Invasive attacks: reverse engineering

- Reverse engineering – understanding the structure of a semiconductor device and its functions
  - optical, using a confocal microscope (for > 0.5 μm chips)
  - deprocessing is necessary for chips with smaller technology

Confocal microscopes represent the different chip layers in different colors. In the right image, the metal interconnects have been removed with hydrofluoric acid. Both images together can be read almost as easily as a circuit diagram.
Invasive attacks: reverse engineering

- Removing top metal layer using wet chemical etching
  - good uniformity over the surface, but works reliably only for chips fabricated with 0.8 μm or larger process (without polished layers)
Invasive attacks: reverse engineering

- Memory extraction from Mask ROMs
  - removing top metal layers for direct optical observation of data in NOR ROMs (bits programmed by presence of transistors)
  - not suitable for VTROM (ion implanted) used in smartcards – selective (dash) etchants are required to expose the ROM bits

NEC μPD78F9116 microcontroller
0.35 μm

Motorola MC68HC05SC27 smartcard
1.0 μm
Picture courtesy of Dr Markus Kuhn
Invasive attacks: microprobing

- Microprobing with fine electrodes
  - eavesdropping on signals inside a chip
  - injection of test signals and observing the reaction
  - can be used for extraction of secret keys and memory contents
  - limited use for 0.35µm and smaller chips
Invasive attacks: microprobing

- Laser cutting systems
  - removing polymer layer from a chip surface
  - local removing of a passivation layer for microprobing attacks
  - cutting metal wires inside a chip
  - maximum can access the second metal layer

Picture courtesy of Dr Markus Kuhn
Invasive attacks: chip modification

- Focused Ion Beam (FIB) workstation
  - chip-level surgery with 10 nm precision
  - etching with high aspect ratio
  - platinum and SiO₂ deposition

Picture courtesy of Semiresearch Ltd
Invasive attacks: chip modification

• Focused Ion Beam workstation
  – creating probing points inside smartcard chips, read the memory
  – modern FIBs allow backside access, but requires special chip preparation techniques to reduce the thickness of silicon
Semi-invasive attacks

• Fill the gap between non-invasive and invasive attacks
  – less damaging to target device (decapsulation without penetration)
  – less expensive and easier to setup and repeat than invasive attacks
  – can overcome many challenges put by security in modern chips

• Tools
  – IC soldering/desoldering station
  – simple chemical lab
  – high-resolution optical microscope
  – UV light sources, lasers
  – oscilloscope, logic analyser, signal generator
  – PC with data acquisition board, FPGA board, prototyping boards
  – special microscopes (laser scanning, infrared etc.)

• Types of semi-invasive attacks: passive and active
  – imaging: optical and laser techniques
  – fault injection: UV attack, photon injection, local heating, masking
  – side-channel attacks: optical emission analysis, induced leakage
Semi-invasive attacks: imaging

- Backside infrared imaging
  - microscopes with IR optics give better quality of image
  - IR-enhanced CCD cameras or special cameras must be used
  - resolution is limited to ~0.6μm by the wavelength of used light
  - view is not obstructed by multiple metal layers
Semi-invasive attacks: imaging

- Backside infrared imaging
  - Mask ROM extraction without chemical etching
- Main option for 0.35\(\mu\)m and smaller chips
  - multiple metal wires do not block the optical path

Texas Instruments MSP430F112 microcontroller
0.35 \(\mu\)m

Motorola MC68HC705P6A microcontroller
1.2 \(\mu\)m
Semi-invasive attacks: imaging

- Advanced imaging techniques – active photon probing (Optical Beam Induced Current (OBIC))
  - photons with energy exceeding semiconductor band gap ionize IC’s regions, which results in a photocurrent flow producing the image
  - used for localisation of active areas
  - also works from the rear side of a chip (using infrared lasers)

Microchip PIC16F84A microcontroller
Semi-invasive attacks: imaging

- Advanced imaging techniques – active photon probing (light-induced voltage alteration (LIVA) technique)
  - photon-induced photocurrent is dependable on the state of a transistor
  - reading logic state of CMOS transistors inside a powered-up chip
  - works from the rear side of a chip (using infrared lasers)
- Requires backside approach for 0.35µm and smaller chips
  - multiple metal wires do not block the optical path
  - resolution is limited to ~0.6µm (still enough for memory cells)

![Sensitivity image](image.png)
Semi-invasive attacks: fault injection

• Optical fault injection attacks
  – optical fault injection was observed in my experiments with microprobing attacks in early 2001, introduced as a new method in 2002
  – lead to new powerful attack techniques and forced chip manufacturers to rethink their design and bring better protection
  – original setup involved optical microscope with a photoflash and Microchip PIC16F84 microcontroller programmed to monitor its SRAM
  – photons can switch MOS transistor to ON state
  – SRAM and flip-flops keep changed state after the light pulse
Semi-invasive attacks: fault injection

• Optical fault injection attacks
  – the chip was decapsulated and placed under a microscope
  – light from the photoflash was shaped with aluminium foil aperture
  – physical location of each memory address by modifying memory contents
  – the setup was later improved with various lasers and a better microscope

• Requires backside approach for 0.35μm and smaller chips
  – successfully tested on chips down to 130nm
Semi-invasive attacks: fault injection

- Localised heating using cw lasers
  - test board with PIC16F628 and PC software for analysis
  - permanent change of a single memory cell on a 0.9µm chip
- Limited influence on modern chips (<0.5µm) – influence on adjacent cells
Semi-invasive attacks: fault injection

- Memory masking attacks
  - temporarily disable write and erase operations in embedded memory (Flash/EEPROM) and write into volatile memory (SRAM)
  - use cw red lasers for front-side and infrared lasers for backside attacks

<table>
<thead>
<tr>
<th>Chip</th>
<th>Memory Write Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Flash Cells</td>
</tr>
<tr>
<td>PIC16F628A</td>
<td>1 – 2</td>
</tr>
<tr>
<td>PIC16F628A (backside)</td>
<td>12 – 45</td>
</tr>
</tbody>
</table>
Semi-invasive attacks: side-channel

• Optically enhanced position-locked power analysis
  – Microchip PIC16F84 microcontroller with test program at 4 MHz
  – classic power analysis setup (10 Ω resistor in GND, digital storage oscilloscope) plus laser microscope scanning setup
  – test pattern
    • run the code inside the microcontroller and store the power trace
    • point the laser at a particular transistor and store the power trace
    • compare two traces
Semi-invasive attacks: side-channel

- Optically enhanced position-locked power analysis
  - results for memory read operations: non-destructive analysis of active memory locations (‘0’ and ‘1’)
  - results for memory write operations: non-destructive analysis of active memory locations (‘0→0’, ‘0→1’, ‘1→0’ and ‘1→1’)
- Only backside approach for 0.35µm and smaller chips
  - single-cell access is limited to 0.5µm laser spot
Semi-invasive attacks: side-channel

- Optical emission analysis
  - transistors emit photons when they switch
  - $10^{-2}$ to $10^{-4}$ photons per switch with peak in NIR region (900–1200 nm)
  - optical emission can be detected with photomultipliers and CCD cameras
  - comes from area close to the drain and primarily from the NMOS transistor
Semi-invasive attacks: side-channel

- Optical emission analysis
  - Microchip PIC16F628 microcontroller with test code at 20 Mhz; PMT vs SPA and CCD camera images in just 10 minutes
- Only backside approach for 0.35µm and smaller chips
  - successfully tested on chips down to 130nm (higher Vcc, >1 hour)
Defence technologies: tamper protection

• Old devices
  – security fuse is placed separately from the memory array (easy to locate and defeat)
  – security fuse is embedded into the program memory (hard to locate and defeat), similar approach is used in many smartcards in the form of password protection and encryption keys
  – moving away from building blocks which are easily identifiable and have easily traceable data paths

Motorola MC68HC908AZ60A microcontroller

Scenix SX28 microcontroller
Defence technologies: tamper protection

- Help came from chip fabrication technology
  - planarisation as a part of modern chip fabrication process (0.5 \( \mu \)m or smaller feature size)
  - glue logic design makes reverse engineering much harder
  - multiple metal layers block any direct access
  - small size of transistors makes attacks less feasible
  - chips operate at higher frequency and consume less power
  - smaller and BGA packages scare off many attackers
Defence technologies: tamper protection

• Additional protections
  – top metal layers with sensors
  – voltage, frequency and temperature sensors
  – memory access protection, crypto-coprocessors
  – internal clocks, power supply pumps
  – asynchronous logic design, symmetric design, dual-rail logic
  – ASICs, secure FPGAs and custom-designed ICs
  – software countermeasures
Defence technologies: what goes wrong

• Security advertising without proof
  – no means of comparing security, lack of independent analysis
  – no guarantee and no responsibility from chip manufacturers
  – wide use of magic words: *protection, encryption, authentication, unique, highly secure, strong defence, cannot be, unbreakable, impossible, uncompromising, buried under x metal layers*

• Constant economics pressure on cost reduction
  – less investment, hence, cheaper solutions and outsourcing
  – security via obscurity approach

• Quicker turnaround
  – less testing, hence, more bugs

• What about back-doors?
  – access to the on-chip data for factory testing purposes
  – how reliably was this feature disabled?
  – how difficult is to attack the access port?
  – are there any trojans deliberately inserted by subcontractors?
Actel/Microsemi ProASIC3 Flash FPGA

- Scanning JTAG for command space
  - find depth of DR registers associated with each command
  - test if those DR registers can be amended

- Analysing STAPL programming file from design software
  - hints on unused spaces
Simplified ProASIC3 security

- AES encryption engine can only send data in one direction
- Passkey only unlocks FROM readback
- Hidden JTAG functions include different areas
  - factory settings, debug features and control registers
  - no references were found in their tools or documentation that readback of the design was a possibility
Defence technologies: how it fails

- Microchip PIC microcontroller: security fuse bug
  - security fuse can be reset without erasing the code/data memory
    - solution: fixed in newer devices
- Hitachi smartcard: information leakage on a products CD
  - full datasheet on a smartcard was placed by mistake on the CD
- Actel secure FPGA: programming software bug
  - devices were always programmed with a 00..00 passkey
    - solution: software update
- Xilinx secure CPLD: programming software bug
  - security fuse incorrectly programmed resulting in no protection
    - solution: software update
- Dallas SHA-1 secure memory: factory initialisation bug
  - some security features were not activated resulting in no protection
    - solution: recall of the batch
- Other possible ways of security failures
  - insiders, datasheets of similar products, development tools, patents
    - solution: test real devices and control the output
Conclusions

• There is no such a thing as absolute protection
  – given enough time and resources any protection can be broken

• Technical progress helps a lot, but has certain limits
  – do not overestimate capabilities of the silicon circuits
  – do not underestimate capabilities of the attackers

• Defence should be adequate to anticipated attacks
  – security hardware engineers must be familiar with attack technologies to develop adequate protection
  – choosing the correct protection saves money in development and manufacturing

• Attack technologies are constantly improving, so should the defence technologies

• Many vulnerabilities were found in various secure chips and more are to be found posing more challenges to hardware security engineers
References

• Slides
  – http://www.cl.cam.ac.uk/~sps32/PartII_030214.pdf

• Literature:
  – http://www.cl.cam.ac.uk/~sps32/#Publications