Fault and side-channel attacks on memory

Dr Sergei Skorobogatov

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

Computer Laboratory

Introduction: Who needs secure chips?

- **car industry:** anti-theft protection, spare parts identification
- **service providers:** access cards, payment tokens, RFID tags, electronic keys, software license dongles
- **mobile phone manufacturers:** batteries and accessories control
- **printer manufacturers:** toner cartridges, memory modules
- **manufacturers of entertainment systems:** copy protection, consumables and accessories control
- **manufacturers of devices and equipment:** protection against cloning and reverse engineering, IP protection (hardware, software, algorithms)
- **banking industry:** secure payment cards, secure processing
- **military applications:** data protection, encrypted communication

Introduction: Why 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

Hardware security evolution

Hardware security evolution

Hardware security

- Embedded memory
	- SRAM, Mask ROM, EEPROM, Flash
	- stores sensitive information, critical parts of algorithms, passwords, encryption keys
	- easy to locate on a die
- Common components
	- CPU
	- Memory
	- $-$ I/O
	- A/D and D/A

Embedded memory

- EEPROM and Flash
	- access one row at a time
	- read-sense amplifiers bottleneck
	- high-voltage operation
- SRAM
	- access with data bus width
	- read-sense amplifiers bottleneck

Choosing secure components

- What has changed in the past?
	- too many 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
	- chip manufacturers will not tell you the truth about security
	- marketing dominates over security

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

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
	- brute forcing
- Compare old days (late 90s) with today challenges

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
- Today: timing attacks became harder to apply
	- common mistakes were fixed by manufacturers
	- internal clock sources and use of PLL made analysis difficult
	- countermeasures are in place: randomised clock, dummy cycles
	- careful selection of hardware eliminates many problems

Non-invasive attacks: side-channel

- 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
- Today: SPA/DPA and EMA became more challenging
	- higher operating frequency and noise: faster equipment is required
	- power supply is reduced from 5V to 1V: lower signal, more noise
	- 8-bit data vs 32-bit data: harder to distinguish single-bit change
	- more complex circuits: higher noise from other parts, hence, more signal averaging and digital signal processing are required
	- effective countermeasures for many cryptographic algorithms

Non-invasive attacks: fault injection

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

cont: … … …

Non-invasive attacks: fault injection

- Today: glitch attacks became harder to exploit
	- effective countermeasures are in place: clock and power supply monitors
	- internal clock sources, clock conditioning and PLL circuits
	- internal charge pumps and voltage regulators
	- asynchronous design
	- checksums (CRC, SHA-1)
	- encryption

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
- Compare old days (late 90s) with today challenges

Invasive attacks: sample preparation

- **Decapsulation**
	- manual with fuming nitric acid (HNO $_{\text{3}}$) and acetone at 60°C
	- automatic using mixture of $HNO₃$ and $H₂SO₄$
	- full or partial
	- from front side and from rear side
- Today: more challenging due to small and BGA packages

Invasive attacks: imaging

- Optical imaging
	- resolution is limited by optics and wavelength of a light:
		- R = 0.61 *λ* / *NA* = 0.61 *λ* / *n* sin(*μ*) best is 0.18µm technology
		- 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 ($n = 1.5$)
- Today: optical imaging is replaced by electron microscopy

Bausch&Lomb MicroZoom, 50×2×, NA = 0.45 Leitz Ergolux AMC, 100×, NA = 0.9

Invasive attacks: reverse engineering

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

Invasive attacks: reverse engineering

- **Deprocessing**
	- removing passivation layer to expose the top metal layer for microprobing attacks
	- decomposition of a chip for reverse engineering
	- Mask ROM extraction
- Methods
	- wet chemical etching (KOH solutions, HCl, ${\sf H_2O_2}$)
		- isotropic uniformity in all directions
		- uneven etching and undercuts metal wires lift off the surface
	- $-$ plasma etching or dry etching (CF $_{4}$, C $_{2}$ F $_{6}$, SF $_{6}$ or CCl $_{4}$ gases)
		- perpendicular to the surface
		- speed varies for different materials
	- $-$ chemical-mechanical polishing (abrasives like Al $_2$ 0 $_3$ or diamond)
		- good planarity and depth control, suitable for modern technologies
		- difficult to maintain planarity of the surface, special tools are required

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 0.8μm or larger process (without polished layers)
- Today: plasma etching and chemical-mechanical polishing

Motorola MC68HC705C9A microcontroller

1.0 μm

NEC μPD78F9116 microcontroller 0.35 μm

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

- Today: Focused Ion Beam workstation
	- chip-level surgery with 10nm precision
	- create probing points inside smartcard chips, read the memory
	- modern FIBs allow backside access, but require special chip preparation techniques to reduce the thickness of silicon

Semi-invasive attacks

- Filling 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
- 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
	- **side-channel attacks:** optical emission analysis, induced leakage
- Compare old days (late 90s) with today challenges

Semi-invasive attacks: sample preparation

- **Decapsulation**
	- manual with fuming nitric acid (HNO $_{\text{3}}$) and acetone at 60°C
	- automatic using mixture of $HNO₃$ and $H₂SO₄$
	- full or partial
	- from front side and from rear side
- Today: more challenging due to small and BGA packages

- 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 \sim 0.6 µm by the wavelength of used light
	- view is not obstructed by multiple metal layers

- Backside infrared imaging – Mask ROM extraction without chemical etching
- Today: the main option for 0.35µm and smaller chips – multiple metal wires do not block the optical path

Texas Instruments MSP430F112 microcontroller

0.35 μm

Motorola MC68HC705P6A microcontroller 1.2 μm

- 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

- Advanced imaging techniques active photon probing
	- light-induced current variation
		- alternative to 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)
- Today: backside approach for 0.35µm and smaller chips
	- multiple metal wires do not block the optical path
	- resolution is limited to $\sim 0.6 \mu m$ (still enough for memory cells)

Microchip PIC16F84 microcontroller

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

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
- Today: 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
- Today: influence is limited for modern chips (<0.5µm)
	- adjacent cells are affected as well

33

- Optically enhanced position-locked power analysis
	- Microchip PIC16F84 microcontroller with test program at 4MHz
	- classic power analysis setup (10Ω 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

- 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 \rightarrow 0', '0 \rightarrow 1', '1 \rightarrow 0' and '1 \rightarrow 1')
- Today: backside approach for 0.35µm and smaller chips

– single-cell access is limited to 0.5µm laser spot

- Optical emission analysis
	- transistors emit photons when they switch
	- − 10⁻² to 10⁻⁴ photons per switch with peak in NIR region (900–1200nm)
	- optical emission can be detected with photomultipliers and CCD cameras
	- comes from area close to the drain and primarily from the NMOS transistor

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

- Optical emission analysis: new challenges
	- Actel® ProASIC3® 0.13μm, 7 metal layers, flash FPGA
	- *"highly secure FPGA"* which is reprogrammable, non-volatile, single-chip and live-at-power-up solution
	- *"offer one of the highest levels of design security in the industry"*
	- robust design security features: flash logic array, flash ROM, security fuses, FlashLock™, AES
	- *"even without any security measures (such as FlashLock with AES), it is not possible to read back the programming data from a programmed device"*
	- allows secure ISP field upgrades using 128-bit AES-encrypted bitstream with AES authentication and MAC verification
	- other security measures: voltage monitors, internal charge pumps, asynchronous internal clock and many others

- Sample preparation of A3P060 FPGA: front and rear
	- the surface is covered with sticky polymer which needs to be removed for physical access to the surface
	- >99% of the surface is covered with supply grid and dummy fillers
	- backside: low-cost approach used without any treatment

- Sample preparation: front
	- only three top metal layers are visible at a most
	- full imaging will require de-layering and scanning electron microscopy
	- any invasive attacks will require sophisticated and expensive equipment

- Backside imaging is the only possibility
	- low spatial resolution of about 1μm (R=0.61λ/NA=0.61·1000/0.5)
- 20× NIR objective lens, light source with Si filter
- Locating internal blocks: JTAG, Flash ROM, SRAM
- Optical emission analysis
	- power supply was increased from 1.5V to 2.0V to boost the emission

- Increasing the power supply voltage: every 10% of increase above nominal Vcc boosts the emission by 40%…120%
- A3P060: JTAG ID reading

- Flash ROM (Settings + Data)
	- 20× NIR objective lens
	- 60' integration time
	- continuous reading
- Recognisable data pattern
	- some data can be extracted
	- gives information about location

- SRAM dedicated for AES
	- 20× NIR objective lens
	- 120' integration time
	- continuous initialisation
- AES key recovery
	- key scheduling used in AES
	- AES key can be easily calculated from any round key
	- existence of separate JTAG commands for AES initialisation, authentication and decryption
	- information is leaked by the SRAM array and write drivers

- SRAM dedicated for AES
	- 20× NIR objective lens
	- 120' integration time
	- continuous initialisation
- **Exploiting power supply trick**
	- alternating the supply voltage during the operation: 2.0V peak
	- 16μs per AES initialisation
	- 1.6μs per each round key: calculation + storage
	- 16 bit at a time: 8 write cycles

- SRAM dedicated for AES
	- 20× NIR objective lens
	- 120' integration time
	- continuous initialisation
- **Exploiting power supply trick**
	- alternating the supply voltage during the last round operation: 2.5V peak
	- 0.2μs increase of the supply voltage from 1.5V to 2.5V for one write cycle

New attacks

- Flash memory 'bumping' attacks
	- will appear at CHES-2010, 17-20 August 2010, Santa Barbara
	- new class of optical fault injection attacks aimed at data and key extraction from embedded memory through indirect access (authentication, verify operation, CRC check, hash check)
- Optical Fault Masking Attacks
	- will appear at FDTC-2010, 21 August 2010, Santa Barbara
	- new types of optical fault attacks aimed at disrupting of the normal memory operation through preventing changes of the memory contents (write and erase protection)

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

Defence technologies: tamper protection

- Help came from chip fabrication technology
	- planarisation as a part of modern chip fabrication process (0.5 μm 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

0.9µm microcontroller 0.5µm microcontroller 0.13µm FPGA

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 the factory testing feature disabled?
	- how difficult is to attack the access port?

Defence technologies: how it fails

- Microchip PIC microcontrollers: security fuse bug *(command)* – security fuse can be reset without erasing the code/data memory
- Atmel AVR microcontrollers: security fuse bug *(glitch attack)* – security fuse can be reset without erasing the code/data memory
- 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
- Xilinx secure CPLD: programming software bug – security fuse incorrectly programmed resulting in no protection
- Maxim/Dallas SHA-1 secure memory: factory setting bug – some security features were not activated resulting in no protection
- Other examples
	- insiders, datasheets of similar products, development tools
	- solution: test real devices and control the output

Defence technologies: why goes wrong?

- Ignorance of mistakes by chip manufacturers
- Unconditional trust from customers
- Reluctance to collaborate with people from academia
- Security perception and awareness levels
	- the security bug cannot be fixed unless the attack procedure is known and reliably reproduced
	- the security flaw cannot be fixed unless the attack method is learned and well understood
- Engineering problem: for many systems the security comes as an extra feature added at a later design stage

Future work

- Improvements to semi-invasive attacks
	- some of 180nm and 130nm chips tested
	- preparation for testing 90nm chips is under way
	- 65nm chips are in plans
- Seeking collaboration with industry
	- evaluation of products against new attacks
	- developing new attack methods and techniques
	- focusing on low-cost attacks which are more dangerous
- New challenges
	- synchronisation techniques for side-channel and fault attacks
	- new imaging techniques based on fault attacks
	- is everything solved in the side-channel attacks area?

New directions for research

- Boosting side-channel attacks with new methods and techniques aimed at improvement by a factor of 1000000
	- off-the-shelf solution vs special hardware
	- what a million times improvement would mean for a real device?
		- 1 day for an attack which normally takes 2000 years to succeed
		- 1 second for an attack which normally takes 10 days to succeed
- Fixed funds and fixed term attacks?
	- how far could an attacker move given X budget and limited time?
- What is 'practical attack'?
	- could someone achieve key extraction within 1 second and 1000\$
- Backdoors testing
	- many chips have Factory test and Debug modes, are they secure?
- Clone dilemma
	- how one can prove that another product is a clone and not a compatible product (forensic analysis within security constraints)?
	- $-$ if a product is cloned, how was it done (there are many ways)? $^{\rm 55}$

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
- 56 • Many vulnerabilities were found in various secure chips and more are to be found, that poses more challenges to hardware security engineers

References

• Abstract

– http://www.cl.cam.ac.uk/~sps32/PASTIS2010_abstract.pdf

• Slides

– http://www.cl.cam.ac.uk/~sps32/PASTIS2010.pdf

- Literature:
	- http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-630.pdf
	- http://www.cl.cam.ac.uk/~sps32/#Publications