Optical surveillance on silicon chips: your crypto keys are visible

Dr Sergei Skorobogatov

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

UNIVERSITY OF CAMBRIDGE
Computer Laboratory
Talk Outline

• Introduction
• Background of optical emission
• Experimental setup
• Results for an old microcontroller chip
• Limitations and improvements
• Challenge with modern chips
• Results for a secure FPGA chip
• Countermeasures
• Conclusion
Introduction

• Operating semiconductor circuits emit photons
  – known for over 40 years
  – actively used in failure analysis for over 20 years

• Existing failure analysis techniques
  – picosecond imaging circuit analysis (PICA) uses photomultiplier array
  – photon emission microscopy (PEM) uses special IR cameras
  – both techniques are expensive and require sophisticated sample preparation

• What about hardware security?
  – any possibility of seeing internal signals?
  – any leaks from memory arrays?
Introduction

• Optical emission analysis attacks were introduced in 2008 and exploit well known fact that photon emission of a chip is correlated with the processed data*
  – done on a PIC16F84A (0.9 μm) running at 6MHz with 7V supply
  – from backside with the silicon substrate thinned down to 20 μm
  – using Mepsicron II camera with hi-res 2D imaging and 50ps timing
  – continued for 12 hours with test code in a loop
  – proved that AES key can be extracted from the operating device

• Can this be used to compromise security in silicon chips?
  – requires expensive equipment and special chip preparation
  – was not considered as a threat, hence, no protection is in place
  – does not form part of standard security evaluation techniques

* J. Ferrigno et al, “When AES blinks: introducing optical side channel”, IET Information Security
Introduction

• Challenges
  – find low-cost detectors suitable for optical emission analysis
  – reduce the cost of sample preparation

• Any technical progress for the past 20 years?
  – are modern CCD cameras good for the attack?
  – what about photomultipliers (PMT)?
  – what parameters are essential for such detectors?

• If optical emission from operating chip has correlation with processed data, is there any correlation between photon emission and power consumption?
  – if found, this can be used for finding weak spots in protection against power analysis attacks
  – optical emission can be scaled down to an individual transistor
Background

• What is the problem with optical emission analysis attacks?
• Number of photons emitted per every switch of a transistor
  \[ N_e = S_e B \left( \frac{L_H I_d}{q v_s} \right) T_s \sim 10^{-2} \ldots 10^{-4} \text{ photons/switch} \]
  \( S_e \) – spectral emission density, \( B \) – emission bandwidth, \( L_H \) – hot-carrier region length,
  \( I_d \) – drain current, \( q \) – e\(^{-}\) charge, \( v_s \) – carrier saturated velocity, \( T_s \) – transition time
• Emission spectrum is from \( \sim 500\text{nm} \) to above \( 1200\text{nm} \) with maximum emission at \( 900\text{nm} \ldots 1100\text{nm} \) (NIR region)
• Small fraction of emitted photons can be detected: <1%
  – emission is isotropic, so with a lens only 25%…45% is observed
  – there are losses in optics due to reflections and absorption (80%)
  – low quantum efficiency (QE) of detectors in NIR region: 1%…20%
• Backside approach: <0.1%
  – high refractive index of silicon \( n_{1000\text{nm}} = 3.58 \) causes high reflection (32%) and low critical angle \( (\theta = 16.2^\circ) \) results in reduced aperture
Background

- Optical emission is higher from the n-MOS transistor due to higher mobility of electrons
- Emission takes place near the drain area where the speed of carriers declines
Experimental setup

• Challenges in choosing the right detector
  – single-photon sensitivity
  – low emission intensity requires longer integration time, hence, detectors must have low noise and low dark current
  – NIR emission spectrum requires detectors sensitive in that area

• Photomultiplier (PMT)
  – single-sensor detector with large aperture
  – fast detection

• Avalanche photodiode (APD)
  – single-sensor detector with small aperture
  – fast detection

• Cameras with charge-coupled devices (CCD)
  – 2D detector with high resolution: 500x500 to 4000x3000
  – very low frame rate: 10μs to 1s
Experimental setup

- Challenges in choosing the right PMT and APD: as good as possible NIR sensitivity, as low as possible dark current
  - PMT usually have very limited NIR sensitivity
  - detectors with better NIR sensitivity have higher dark current
  - low dark current in APD is caused by their small aperture size
  - too small aperture size of APD (10μm…500μm) complicates their usage

<table>
<thead>
<tr>
<th>Type of detector</th>
<th>Wavelength, nm</th>
<th>QE at 900nm</th>
<th>QE at 1000nm</th>
<th>Dark current, e⁻/s</th>
<th>Time response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantar Mepsicron II, S25</td>
<td>180–940</td>
<td>1%</td>
<td>0%</td>
<td>0.005</td>
<td>50ps</td>
</tr>
<tr>
<td>Hamamatsu H10330-25</td>
<td>850–1250</td>
<td>2%</td>
<td>2%</td>
<td>2000</td>
<td>900ps</td>
</tr>
<tr>
<td>Hamamatsu H6780-01</td>
<td>250–850</td>
<td>0%</td>
<td>0%</td>
<td>400</td>
<td>780ps</td>
</tr>
<tr>
<td>Sensl PCDMini-0020</td>
<td>400–1100</td>
<td>2%</td>
<td>1%</td>
<td>50</td>
<td>200ps</td>
</tr>
</tbody>
</table>
Experimental setup

- Challenges in choosing the right CCD camera: as good as possible NIR sensitivity, as low as possible dark current
  - monochrome cameras have good NIR sensitivity
  - CCTV and hobbyist astronomical cameras have low dark current and good NIR sensitivity

<table>
<thead>
<tr>
<th>Type of detector</th>
<th>Wavelength, nm</th>
<th>QE at 900nm</th>
<th>QE at 1000nm</th>
<th>Dark current, e⁻/s</th>
<th>Time response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantar Mepsicron II, S25</td>
<td>180–940</td>
<td>1%</td>
<td>0%</td>
<td>0.005</td>
<td>50ps</td>
</tr>
<tr>
<td>Hamamatsu C4880-21</td>
<td>200–1200</td>
<td>50%</td>
<td>20%</td>
<td>0.3</td>
<td>20ms</td>
</tr>
<tr>
<td>Hamamatsu C4880-50</td>
<td>200–1100</td>
<td>30%</td>
<td>10%</td>
<td>0.01</td>
<td>20ms</td>
</tr>
<tr>
<td>Average monochrome CCD</td>
<td>400–1000</td>
<td>5%</td>
<td>1%</td>
<td>1</td>
<td>20ms</td>
</tr>
<tr>
<td>Average colour CCD</td>
<td>400–700</td>
<td>0%</td>
<td>0%</td>
<td>1</td>
<td>20ms</td>
</tr>
<tr>
<td>Sony Super HAD CCD</td>
<td>300–1050</td>
<td>8%</td>
<td>1%</td>
<td>0.02</td>
<td>10μs</td>
</tr>
<tr>
<td>Sony EXview HAD CCD</td>
<td>300–1100</td>
<td>12%</td>
<td>5%</td>
<td>0.02</td>
<td>10μs</td>
</tr>
</tbody>
</table>
Experimental setup

- Sample preparation: PIC16F628 microcontroller (0.9μm)
- Locating internal blocks: Flash, EEPROM, SRAM, CPU
- Running the chip at 20MHz clock (5MIPS) with 6V power supply to boost the emission
Experimental setup

• PMT setup: decapsulated chip facing sensor's aperture
  – Hamamatsu H6780-01 PMT sensor

• CCD setup: camera mounted on a microscope with the chip placed in a test socket
  – Starlight Xpress SXV-H9 CCD camera
Results

- **PMT**: 60' acquisition time, digital storage oscilloscope in color-graded mode with infinite persistence with histogram
- **SPA**: 10Ω resistor, digital storage oscilloscope with active probe
- **Test code**: `bsf portb,3
cclf 0x75
decl 0x75,f
dcf portb,3
goto loop`

- **PMT vs SPA**
  - higher bandwidth
  - special hardware will suit better as oscilloscope is not designed for long-time integration (latency issue)
Results

• PMT vs SPA
  – higher bandwidth provides more data for analysis
  – possible localisation of source through apertures and optics
  – good correlation suggests possibility of using optical emission analysis for characterisation of areas contributing to power trace
  – acquisition of emission requires some time with the device under test performing the same operation and precisely synchronised
Results

- **CCD**
  - 2× objective lens
  - 30' integration time
  - EEPROM data: 00h, FFh
  - SRAM data: variable 00h…FFh
  - continuous EEPROM reading and SRAM writing and reading

- **Test code:**
  - incf EEADR,f
  - bsf EECON1,RD
  - movf EEDATA,w
  - decf 0x75,f
  - goto loop

- **2D image with recognisable areas of emission from Flash, EEPROM, SRAM and CPU**
Results

• CCD with high magnification
  – 100× objective lens
  – 10' integration time
  – EEPROM data: 00h, FFh
  – continuous EEPROM reading

• Test code:
  ```
  incf EEADR,f
  bsf EECON1, RD
  movf EEDATA, w
  goto loop
  ```

• Emission from the NMOS transistor is significantly higher than from the PMOS
Results

• EEPROM area
  – 10× objective lens
  – 10' integration time
  – data: 56h, 56h, 56h…56h, 00h
  – continuous EEPROM reading

• Test code: in cf EEADR,f
  bsf EECON1,RD
  movf EEDATA,w
  goto loop

• Flash memory has similar structure and gives similar result
  – data extraction is complicated by the fact that program code is executed from the flash memory
Results

• SRAM area
  – 10× objective lens
  – 10' integration time
  – data: A6h, W=A6h
  – continuous reading and writing

• Test code: \texttt{movf 0x75, w} \texttt{movwf 0x75}
  \texttt{goto loop} \texttt{goto loop}

• Low emission from memory cells
  – write drivers, bus drivers, row and column selectors leak the most

• Write data have the same emission for ‘0’ and ‘1’
  – dual-rail logic used in SRAM: separate bit lines for writing ‘0’ & ‘1’
  – difference in the emission could predict leakage in the power trace
Results

• **SRAM area**
  - 10× objective lens
  - 10' integration time
  - data: A6h, W=C3h
  - continuous XOR operation

• **Test code:**
  
  ```
  movlw 0xA6
  movwf 0x74
  movlw 0xC3
  xorwf 0x74, f
  goto loop
  ```

• **Leakage through both read and write logic**
  - read: intensity is proportional to the number of ‘1’s
  - write: ‘0’s and ‘1’s are separated
Limitations and improvements

• Data recovery
  – slow process: minimum 1 minute per byte

• Modern chips
  – three or more metal layers prevent direct observation and analysis
  – smaller technologies will require longer integration time

• Backside approach
  – silicon is transparent to light with wavelengths above 1000 nm
  – lower spatial resolution of ~1μm (R=0.61λ/NA)
  – longer integration time due to higher losses in silicon and optics
  – higher magnification lenses give better result
  – use of NIR optics improves result, but expensive
  – substrate thinning and AR coating are useful, but expensive
  – increase of the power supply voltage boosts the optical emission
Limitations and improvements

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

- PIC16F628: EEPROM reading

<table>
<thead>
<tr>
<th>Power supply voltage</th>
<th>3.5V</th>
<th>4.0V</th>
<th>4.5V</th>
<th>5.0V</th>
<th>5.5V</th>
<th>6.0V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photometry results</td>
<td>1046</td>
<td>1286</td>
<td>2427</td>
<td>8400</td>
<td>23292</td>
<td>43026</td>
</tr>
</tbody>
</table>
Limitations and improvements

- PIC16F628 (0.9μm) vs PIC16F628A (0.5μm):
  higher density with CMP technology: approx. 5 times lower intensity
Limitations and improvements

- PIC16F628: EEPROM area from front and rear sides
  higher reflections and absorption in Si: approx. 10 times lower intensity
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
  - “unique in being highly resistant to both invasive and noninvasive attacks”
Experimental setup

• 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 or dummy fillers
  – backside: low-cost approach used – without any treatment
Experimental setup

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

• 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
Experimental setup

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

<table>
<thead>
<tr>
<th>Power supply voltage</th>
<th>1.5V</th>
<th>1.6V</th>
<th>1.8V</th>
<th>2.0V</th>
<th>2.2V</th>
<th>2.5V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photometry results</td>
<td>889</td>
<td>1194</td>
<td>1953</td>
<td>5270</td>
<td>9536</td>
<td>23270</td>
</tr>
</tbody>
</table>
Results

• JTAG glue logic
  – 20× NIR objective lens, 60' integration time
  – repeating the same operation
• Some recognisable differences
• Partial reverse engineering – information
  – operation-related activity
  – obfuscated data flow paths
  – security-related operations
Results

• 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
Results

• 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
Results

• 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
Results

• 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
Countermeasures

• Use of modern chips with multiple metal layers forces an attacker to use backside approach and results in longer time required for the attack
• Metal shielding over sensitive areas can help but cannot prevent backside analysis
• Adding dummy cycles to normal operations
• Encryption makes analysis harder
• Asynchronous circuits could make the attack more problematic as data analysis requires synchronisation
Conclusion

• Optical emission analysis can be carried out at a relatively low cost using hobbyist astronomical CCD cameras with low-magnification optics
• Long exposure time is required: the device must perform the same operation millions of times in a loop
• PMT offers high bandwidth and acquired data have correlation with power analysis results and can be used for finding weak spots in protection against power analysis attacks
• Optical emission analysis offers possibility for partial reverse engineering of chips including data analysis
• Backside approach can help in modern chips, but has lower spatial resolution and requires longer integration time
• Increase of the power supply voltage boosts the optical emission and considerably reduces the time of analysis
• Modern deep-submicron chips do leak information through optical emission when their power supply is increased by at least 30%
• Lack of protection against optical side-channel attacks in modern chips might lead to possible vulnerabilities
Further reading

- Up-to-date information: http://www.cl.cam.ac.uk/~sps32/