CSE 509: System Security
Spring 2019
Side-Channels
Amir Rahmati
Side Channel

- Any data channel that was not part of the regular I/O of the system.
- Examples:
  - Computation time
  - Power Consumption
  - EM radiations
  - Sound
Side Channel Attacks

- Use of side-channels to find secrets about the system.
- Usually happens in two phases:
  - Monitoring phase:
    - Collect data during normal operation
  - Analysis phase:
    - Find patterns in data
Simple Example: Password Checker
Timing Attacks

For c1 in password_1 and c2 in password_2:
    if c1 == c2:
        continue
    else:
        password check failure

password check success
Encryption
Encryption

- Find the key
- Find the intermediary values —> Weaken the key
Simple Power Analysis

- Distinguish 1s from 0s
- Distinguish addition from multiplication
- Ratios of 0’s and 1’s
Differential Power Analysis

- Statistical Analysis
Example: AES Algorithm

- 128-bit length key $\rightarrow 2^{128}$ guesses
- 128-bit AES can be broken into 16 byte-chunks
- Testing each byte individually (256 options)
Van Eck Phreaking

- Electrons moving in a conductor emit EM waves
- Put Antenna
- Run side channel analysis

https://www.youtube.com/watch?v=C7yQ_Z4vB4c
Acoustic Cryptanalysis

- Printer
- Keyboard
- Phone
- Coil-Whine
Side-Channels are not limited to old systems
Precise computation is **not** required in many applications: Machine learning, sensory data, information retrieval, physical simulation, computer vision…
Privacy Implications of Approximate DRAM

Identify the origin of data by looking at the error pattern
Overview
Background on DRAM

- Cell value
- Charge leakage
- Refresh
Toy Example

Device A

Device B

Amir Rahmati

Deanonymizing Approximate Memory
### Distance Metric

<table>
<thead>
<tr>
<th>Bits</th>
<th>A</th>
<th>B</th>
<th>A+</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

#### Hamming Distance

- A and B: 2
- A+ and C: 3
- A and C: 3

#### Jaccard Distance

- A and B: 1
- A+ and C: 1
- A and A+: 1

#### Jaccard Distance (Hamming Weight)

- A and B: 0.5
- A+ and C: 0
- A and A+: 0
- A and C: 0.33
Putting Memory Fingerprint Together

Page granularity

Diagram:

- D1
- D2
- ... (ellipses)
- Dn

Variable:

- l1
- l2
- ... (ellipses)
- ln
Experimental Setup
Uniqueness

How unique are the fingerprints?

Two order of magnitude difference
Order of Failure

Does the fingerprint hold across different levels of approximation?
Level of Approximation

How do different levels of approximation affect identification?
Thermal Effect

How does change in temperature affect identification?

![Graph showing the effect of temperature on identification accuracy](image)
Consistency

How consistent are the fingerprints?
Types of Attack
End to End Feasibility

- Commodity system
- Edge detection tool
- 1000X10MB traces
Chance of Mismatch

How much entropy does a page of memory provide?

For memory of size $M$ bits where $A$ bits of errors are tolerated:

$$\text{Max unique fingerprints} = \binom{M}{A}$$

Given noise threshold of $T$ bits using Hamming bound:

$$\frac{\sum_{i=1}^{T} \binom{M}{i}}{\binom{M}{A}} \leq \text{Chance of mismatching} \leq \frac{\sum_{i=1}^{2T} \binom{M}{i}}{\binom{M}{A}}$$

<table>
<thead>
<tr>
<th>One page of memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M = 32768$ bits, $A = 1%$, $T = 32$ bits</td>
</tr>
</tbody>
</table>

| Max possible fingerprints | $8.70 \times 10^{795}$ |
| Max unique fingerprints   | $\geq 1.07 \times 10^{590}$ |
| Chance of mismatching    | $\leq 9.29 \times 10^{-591}$ |
| Total Entropy            | 2423 bits |

Amir Rahmati

Deanonymizing Approximate Memory
Consider **Security & Privacy** as a primary design criteria in emerging systems

https://github.com/impedimentToProgress/ProbableCause
Can side-channels be beneficial?
Can side-channels be beneficial?
Case Study: Medical Devices
Recently, the compounder was infected with a virus. It is unknown what effect this virus should have on the operation of the software.
<table>
<thead>
<tr>
<th></th>
<th>No software changes</th>
<th>No updates</th>
<th>No manual configuration</th>
<th>No network connection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antivirus</strong></td>
<td>☒</td>
<td>☒</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Firewall</strong></td>
<td>✓</td>
<td>✓</td>
<td>☒</td>
<td>☒</td>
</tr>
<tr>
<td><strong>NIDS</strong></td>
<td>✓</td>
<td>☒</td>
<td>☒</td>
<td>☒</td>
</tr>
<tr>
<td><strong>WattsUpDoc</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Our Insight

• Embedded medical devices have specific, limited tasks

• If we can characterize those tasks, we can identify later deviations
WattsUpDoc
Power Analysis

![Current vs Time Graph]
## Devices Tested

<table>
<thead>
<tr>
<th>Device</th>
<th>Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baxa ExactaMix 2400 compounder</td>
<td>WinXP Embedded, Via 664 MHz, 512 MB RAM</td>
</tr>
<tr>
<td>Schweitzer SEL3354 substation computer</td>
<td>WinXP Embedded, Athlon 2600+, 2 GB RAM</td>
</tr>
</tbody>
</table>
Trace Collection

Measurement points

Sense resistor (behind outlet)

Storage
Compounder Traces

Idle

Infected
Building a Classifier

<table>
<thead>
<tr>
<th>State</th>
<th>Trace</th>
<th>Feature extraction</th>
<th>Label</th>
<th>Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flushing</td>
<td>x,y,z,...</td>
<td></td>
<td>Clean</td>
<td>Watts UpDoc</td>
</tr>
<tr>
<td>Malware</td>
<td>x,y,z,...</td>
<td></td>
<td>Infected</td>
<td>Watts UpDoc</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Malware</th>
<th>x,y,z,...</th>
<th></th>
<th>Watts UpDoc</th>
<th>Clean</th>
</tr>
</thead>
</table>

Frequency (Hz) | Power (x,y,z,...)
---|---
Metrics

Accuracy: \[ \frac{\text{True positives} + \text{True Negative positives}}{\text{True positives} + \text{negatives}} \]

Precision: \[ \frac{\text{True positives}}{\text{True positives} + \text{False positives}} \]

Recall: \[ \frac{\text{True positives}}{\text{True positives} + \text{False negatives}} \]
## Results

<table>
<thead>
<tr>
<th>Device</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compounder</td>
<td>94%</td>
<td>93.7%</td>
<td>76%</td>
</tr>
<tr>
<td>Substation Computer</td>
<td>99.5%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>
# Requirements

<table>
<thead>
<tr>
<th></th>
<th>No software changes</th>
<th>No updates</th>
<th>No manual configuration</th>
<th>No network connection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antivirus</strong></td>
<td>X</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Firewall</strong></td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>NIDS</strong></td>
<td>✓</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>WattsUpDoc</strong></td>
<td>✓</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Testing Unknown Malware

Training

- Sality
- Ramnit
- Autorun
- Bamital

Testing

- Dorkbot
- Bredolab
- Delf
## Results of Unknown Malware

<table>
<thead>
<tr>
<th>Device</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compounder</td>
<td>88.5%</td>
<td>93.5%</td>
<td>92.1%</td>
</tr>
<tr>
<td>Substation Computer</td>
<td>84.9%</td>
<td>98.3%</td>
<td>80.8%</td>
</tr>
<tr>
<td>Behavior-based AV</td>
<td>86%</td>
<td>100%</td>
<td>86%</td>
</tr>
</tbody>
</table>

[Fredrikson et al., IEEE S&P 2010]
Can side-channels be beneficial?  
Case Study: Embedded Devices
Batteryless Devices

**Transportation**

**Passports**

**RFID Sensors**

**Employee IDs**

**Things in Common**
- No long running clocks
- Adversary controls power & time
- Hold secrets
Threats

- Power Analysis
- Reverse Engineering
- Brute Force

Semi-invasive
# Vulnerable to Brute Force Attacks

<table>
<thead>
<tr>
<th>Device</th>
<th>#Queries</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>UHF RFID Tags [Shamir’07]</td>
<td>200</td>
<td>2 Seconds</td>
</tr>
<tr>
<td>MIFARE Classic [Garcia’09]</td>
<td>1,500</td>
<td>16 Seconds</td>
</tr>
<tr>
<td>Digital Signal Transponder [Bono’05]</td>
<td>75,000</td>
<td>1 Hour</td>
</tr>
<tr>
<td>MIFARE DESFire [Paar’11]</td>
<td>250,000</td>
<td>7 Hours</td>
</tr>
<tr>
<td>GSM SIM Cards [Goldberg’99]</td>
<td>150,000</td>
<td>8 Hours</td>
</tr>
</tbody>
</table>
A Time-Keeping Technique Based on SRAM Decay
SRAM Remanence

PMOS Transistors

A ≈ V_{CC}

B ≈ 0V

NMOS Transistors

GND
SRAM Remanence
SRAM Remanence

Seconds

150 190 210
The TARDIS Algorithm

- Initialize SRAM
- SRAM cells decay
- Compute SRAM decay

Voltage

power-up
power-off
power-up

Temperature Later...
Factors Influencing SRAM Decay

- Circuit Capacitance
- Temperature
- Chip Variation
- SRAM Size
Experimental Setup

- DAQ
- TI MSP430
- Thermometer
- Thermal Chamber
TARDIS tells that less than 175s have elapsed since power down.

TARDIS estimates elapsed time.

TARDIS tells that greater than 225s have elapsed since power down.

Stage 1
Stage 2
Stage 3

Seconds without Power

% Memory Decay

26°C, 10μF
## Circuit Capacitance

<table>
<thead>
<tr>
<th>Capacitor Size</th>
<th>Expiration time</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>~0μF</td>
<td>$2.1 \times 10^0$ s</td>
<td>Seconds</td>
</tr>
<tr>
<td>10μF</td>
<td>$2.25 \times 10^2$ s</td>
<td>Minutes</td>
</tr>
<tr>
<td>100μF</td>
<td>$1.98 \times 10^3$ s</td>
<td>1/2 Hour</td>
</tr>
<tr>
<td>1000μF</td>
<td>$2.12 \times 10^4$ s</td>
<td>Hours</td>
</tr>
<tr>
<td>10000μF</td>
<td>$&gt;1.96 \times 10^5$ s</td>
<td>Days</td>
</tr>
</tbody>
</table>

**Batteryless Sensor = 100,000μF**
The nominal current consumed in this mode is only the voltage reached during a power-down. The bottom graph shows the voltage to decay quickly while the microcontroller is active. With so little current being consumed, the supply voltage decays very slowly.

Impact of Temperature:

Increasing the temperature leads to more rapid memory decay for two reasons. First, increasing the temperature increases the leakage currents leaving it floating, reducing voltage and memory decay times by at least an order of magnitude. Second, the decay curves predicted by the 3-parameter models are consistent with the observed decay datasets. The fit lines in the decay plot are that predict decay curves in best agreement with the observed data, as consistent with the empirical explicit capacitance.

The observed current is between 0.05 µA and 2 µA for the 256 KB, 2 KB, and 8 KB of SRAM respectively have 8 KB, 2 KB, and 256 bytes of SRAM. For the MSP430F2131 of the MSP(-:F; family with different SRAM sizes exhibit different decay times but follow the same general trend. For example, Figure 7 shows the 3-parameter DRV probabilities of Equation (3) for different temperatures and capacitance. The DRV probability for the 256 KB, 2 KB, and 8 KB of SRAM at 50°C, 0 µF is almost zero, while the DRV probability for 28°C, 0 µF is close to 100%.

Figure 8: Regardless of temperature and capacitance, the total current comprises the operating current of the microcontroller and the SRAM's data-retention current; both currents are functional explicit capacitance. Implementing the standard decoupling capacitors with additional capacitance reduces the duration of both stage 1 and 2 decay by approximately 80% for the 256 KB, 2 KB, and 8 KB of SRAM. This current is negligible when the microcontroller is active. With so little current being consumed, the supply voltage decays very slowly. This current is between 0.05 µA and 2 µA for the 256 KB, 2 KB, and 8 KB of SRAM respectively have 8 KB, 2 KB, and 256 bytes of SRAM. For the MSP430F2131 of the MSP(-:F; family with different SRAM sizes exhibit different decay times but follow the same general trend. For example, Figure 7 shows the 3-parameter DRV probabilities of Equation (3) for different temperatures and capacitance. The DRV probability for the 256 KB, 2 KB, and 8 KB of SRAM at 50°C, 0 µF is almost zero, while the DRV probability for 28°C, 0 µF is close to 100%.
Section 10.1

The nominal current consumed in this mode is only the voltage reached during a power-down. The bottom graph bytes of SRAM in the MSP(-:F;,-, [data-retention current, specified to be RAM-retention mode in an attempt to avoid losing data. deactivates and kills all clocks to enter an ultra-low power remains active. the voltage to decay quickly while the microcontroller ble by comparison. The high current consumption causes of the MSP430F2131 [the lowest-power operating point (1.8V with 1MHz clock) 350

ational0 The observed current is between ages are above
decay is shown in Fig. 14% and explained here:
tions of the supply voltage. The current during the voltage SRAM's data9retention current; both currents are func9
the operating current of the microcontroller and the
Low Leakage Current:
plementing the standard decoupling capacitors with addi-
the decay curves predicted by the -9parameter models that predict decay curves in best agreement with the ob9
decay depends almost entirely on the minimum supply Figure 8: Regardless of temperature and capacitance, the
A
\[\text{DECAY CURVES} \]
\[
\begin{align*}
\text{Temperature} & \quad 50^\circ \text{C}, 0\mu\text{F} \\
40^\circ \text{C}, 0\mu\text{F} & \\
39^\circ \text{C}, 0\mu\text{F} & \\
32^\circ \text{C}, 0\mu\text{F} & \\
28^\circ \text{C}, 0\mu\text{F} & 
\end{align*}
\]
\[
\begin{align*}
\text{Seconds without Power} & \\
0 & \\
1 & \\
2 & \\
3 & \\
4 & \\
5 & 
\end{align*}
\]
\[
\begin{align*}
\text{% Memory Decay} & \\
0 & \\
10 & \\
20 & \\
30 & \\
40 & \\
50 & 
\end{align*}
\]
Compensate for Temperature

Voltage

power-up

power-off

- Compute SRAM decay
- Analyze Temperature
- Initialize SRAM
Implementation

UHF computational RFID tags augmented with piezo elements

Expiration = 12s
\( \sigma = 0.11s \)
### The Effect of TARDIS*

<table>
<thead>
<tr>
<th>Device</th>
<th>#Queries</th>
<th>W/O TARDIS</th>
<th>W/ TARDIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>UHF RFID Tags</td>
<td>200</td>
<td>2 Seconds</td>
<td>40 Minutes</td>
</tr>
<tr>
<td>MIFARE Classic</td>
<td>1,500</td>
<td>16 Seconds</td>
<td>5 Hours</td>
</tr>
<tr>
<td>Digital Signal Transponder</td>
<td>75,000</td>
<td>1 Hour</td>
<td>10 Day</td>
</tr>
<tr>
<td>MIFARE DESFire</td>
<td>250,000</td>
<td>7 Hours</td>
<td>35 Days</td>
</tr>
<tr>
<td>GSM SIM Cards</td>
<td>150,000</td>
<td>8 Hours</td>
<td>21 Days</td>
</tr>
</tbody>
</table>

* Assuming a 12 seconds TARDIS
Summary

• Provide a notion of time to transiently-powered embedded systems.

• Empirically evaluated the behavior of SRAM-based timekeeping.

• Implemented TARDIS along with three representative applications on an RFID powered microcontroller.