Multi-Tuple Leakage Detection and the Dependent Signal Issue

Olivier Bronchain    Tobias Schneider    François-Xavier Standaert

CHES 2019, Atlanta, USA
Table of Contents

Introduction

Leakage Detection

Multi-Tuple Leakage Detection

Conclusion
Content

Introduction

Leakage Detection

Multi-Tuple Leakage Detection

Conclusion
Side-Channel Issue

Encryption on physical devices:
- Standard utilization
Side-Channel Issue

Encryption on physical devices:
- Standard utilization
- But with any physical signals
Side-Channel Issue

Encryption on physical devices:

- Standard utilization
- But with any physical signals
- Possibly containing secret information
Side-Channel Issue

Encryption on physical devices:
- Standard utilization
- But with any physical signals
- Possibly containing secret information

Side-channel Attacks:
- Known to be hard to prevent
- Hard to evaluate as well
Encryption on physical devices:
- Standard utilization
- But with any physical signals
- Possibly containing secret information

Side-channel Attacks:
- Known to be hard to prevent
- Hard to evaluate as well

Two evaluation approaches:
- Attack based
- Leakage detection
Can directly mount attacks:

1. Collect measurements
Attack Based Evaluation

Can directly mount attacks:
1. Collect measurements
2. Perform an attack
Can directly mount attacks:
1. Collect measurements
2. Perform an attack
3. Retrieve the correct sub-key
Attack Based Evaluation

Can directly mount attacks:
1. Collect measurements
2. Perform an attack
3. Retrieve the correct sub-key

This requires:
1. Long measurement period
2. Skilled/expert knowledge
3. Distinguish 1 sub-key within 256
Leakage Detection Based Evaluation

Leakage detection searches for dependency between manipulated data and physical traces.
Leakage Detection Based Evaluation

*Leakage detection searches for dependency between manipulated data and physical traces.*

- Feed the core with two different sets of inputs
Leakage Detection Based Evaluation

*Leakage detection searches for dependency between manipulated data and physical traces.*

- Feed the core with two different sets of inputs
- Record the corresponding traces
Leakage Detection Based Evaluation

Leakage detection searches for dependency between manipulated data and physical traces.

- Feed the core with two different sets of inputs
- Record the corresponding traces
- Observe differences between the two sets
Leakage Detection Based Evaluation

*Leakage detection searches for dependency between manipulated data and physical traces.*

How does it compare with attack based evaluations:
- Shortened measurement period (Possibly)
- No skilled/expert knowledge
Leakage Detection Based Evaluation

Leakage detection searches for dependency between manipulated data and physical traces.

How does it compare with attack based evaluations:

- Shortened measurement period (Possibly)
- No skilled/expert knowledge

A good first check but:

- Risk of false positives and false negatives
Content

Introduction

Leakage Detection

Multi-Tuple Leakage Detection

Conclusion
Leakage Detection

Find a difference between the two sets:
Leakage Detection

Find a difference between the two sets:

1. Select a point in time
Leakage Detection

Find a difference between the two sets:

1. Select a point in time
2. Record traces to observe a distribution

Distributions

- No difference found
Find a difference between the two sets:

1. Select a point in time
2. Record traces to observe a distribution
3. Perform a statistical test
Leakage Detection

Find a difference between the two sets:
1. Select a point in time
2. Record traces to observe a distribution
3. Perform a statistical test
4. Observe its binary output

The statistical test can search for differences in:
- Means with the Welch’s $t$-test
- Distributions with $\chi^2$-test

No difference found
Leakage Detection

Find a difference between the two sets:
1. Select a point in time
2. Record traces to observe a distribution
3. Perform a statistical test
4. Observe its binary output
Repeat with more measurements if needed

The statistical test can search for difference in:

- Means with the Welch’s $t$-test
- Distributions with $\chi^2$-test

No difference found
Leakage Detection

Find a difference between the two sets:
1. Select a point in time
2. Record traces to observe a distribution
3. Perform a statistical test
4. Observe its binary output

Repeat with more measurements if needed

The statistical test can search for differences in:
- Means with the Welch’s $t$-test
- Distributions with $\chi^2$-test
- . . .
Leakage Detection

Find a difference between the two sets:
1. Select a point in time
2. Record traces to observe a distribution
3. Perform a statistical test
4. Observe its binary output

Repeat with more measurements if needed

The statistical test can search for difference in:
- Means with the Welch’s t-test
- Distributions with $\chi^2$-test
- . . .
Leakage Detection

Find a difference between the two sets:
1. Select a point in time
2. Record traces to observe a distribution
3. Perform a statistical test
4. Observe its binary output
Repeat with more measurements if needed

The statistical test can search for difference in:
- Means with the Welch’s $t$-test $\Rightarrow$ Most popular
- Distributions with $\chi^2$-test
- ...
Leakage Detection: TVLA

The traces contain multiple points in time:

1. Select **all** the points in time
Leakage Detection: TVLA

The traces contain multiple points in time:

1. Select **all** the points in time
2. Record traces to observe a distribution
Leakage Detection: TVLA

The traces contain multiple points in time:

1. Select all the points in time
2. Record traces to observe a distribution
3. Perform independent statistical test

No difference found if:

- At least one of the tests goes above a threshold

Selected thanks to:

- Desired confidence
- Number of considered time samples
- Assuming independence between them
Leakage Detection: TVLA

The traces contain multiple points in time:
1. Select **all** the points in time
2. Record traces to observe a distribution
3. Perform **independent** statistical test
4. Observe their binary outputs

Difference found if:
- At least one of the tests goes above a threshold
- Selected thanks to:
  - Desired confidence
  - Number of considered time samples
  - Assuming independence between them
Leakage Detection: TVLA

The traces contain multiple points in time:
1. Select all the points in time
2. Record traces to observe a distribution
3. Perform independent statistical test
4. Observe their binary outputs
Leakage Detection: TVLA

The traces contain multiple points in time:
1. Select all the points in time
2. Record traces to observe a distribution
3. Perform independent statistical test
4. Observe their binary outputs

Difference found if:
- At least one of the tests goes above a threshold
- Selected thanks to:
  - Desired confidence
  - Number of considered time samples
  - Assuming independence between them
Leakage Detection: TVLA

The traces contain multiple points in time:

1. Select **all** the points in time
2. Record traces to observe a distribution
3. Perform **independent** statistical test
4. Observe their binary outputs

Difference found if:

- At least one of the tests goes above a threshold
Leakage Detection: TVLA

The traces contain multiple points in time:
1. Select all the points in time
2. Record traces to observe a distribution
3. Perform independent statistical test
4. Observe their binary outputs

Difference found if:
- At least one of the tests goes above a threshold
- Selected thanks to:
  - Desired confidence
  - Number of considered time samples
  - Assuming independence between them
Limitations to TVLA

TVLA performs independent $t$-test:

- Impossible to take advantage of multivariate leakage
- Could lead to reduced measurement period
- Independence in the signal is usually not met:
  - Wrong assumption while setting the threshold
  - Hard to interpret results (especially negative ones)
Limitations to TVLA

TVLA performs independent $t$-test:

- Impossible to take advantage of multivariate leakage
Limitations to TVLA

TVLA performs independent $t$-test:
- Impossible to take advantage of multivariate leakage
- Could lead to reduced measurement period
Limitations to TVLA

TVLA performs independent $t$-test:
- Impossible to take advantage of multivariate leakage
- Could lead to reduced measurement period

Independence in the signal is usually not met:
Limitations to TVLA

TVLA performs independent $t$-test:
- Impossible to take advantage of multivariate leakage
- Could lead to reduced measurement period

Independence in the signal is usually not met:
- Wrong assumption while setting the threshold
Limitations to TVLA

TVLA performs independent $t$-test:
- Impossible to take advantage of multivariate leakage
- Could lead to reduced measurement period

Independence in the signal is usually not met:
- Wrong assumption while setting the threshold
  - Hard to interpret results (especially negative ones)
Content

Introduction

Leakage Detection

Multi-Tuple Leakage Detection

Conclusion
Multi-Tuple Leakage Detection: General Idea

Approach:

- **t-test**
  - Replace the independent tests by a single one
  - Natural candidate: Hotelling’s $T^2$-test
  - Do not assume independence
  - Need to invert a covariance matrix
  - Not always applicable

- **D-test**
  - Assume independence
  - Hard to interpret results
Multi-Tuple Leakage Detection: General Idea

Approach:
- Replace the independent tests by a single one
Multi-Tuple Leakage Detection: General Idea

Approach:
- Replace the independent tests by a single one

Multivariate statistical test

Single binary output
Multi-Tuple Leakage Detection: General Idea

Approach:
- Replace the independent tests by a single one

**Natural candidate: Hotelling’s $T^2$-test**
- Do not assume independence
- Need to invert a covariance matrix
  - Not always applicable
Multi-Tuple Leakage Detection: General Idea

Approach:
- Replace the independent tests by a single one

**Natural candidate: Hotelling’s $T^2$-test**
- Do not assume independence
- Need to invert a covariance matrix
  - Not always applicable

**Heuristic alternative: $D$-test**
- Assume independence
  - Hard to interpret results
Traces Parameter: Density

**Density** of informative points:
- The proportion of leaking points
- $t$-test showing difference with $\infty$ of measurements
**Traces Parameter: Density**

**Density** of informative points:

- The proportion of leaking points
- $t$-test showing difference with $\infty$ of measurements

Typical settings:

- Protected software: low density, long traces
- Hardware unprotected: high density, short traces

$\begin{align*}
\text{Density} = 0.1 \\
\text{Density} = 0.2 \\
\text{Density} = 0.5 \\
\text{Density} = 0.9
\end{align*}$
**Traces Parameter: Density**

**Density** of informative points:

- The proportion of leaking points
- $t$-test showing difference with $\infty$ of measurements

![Diagram showing traces with varying density](image)

- $t$-test
- Density = 0.2
Traces Parameter: Density

**Density** of informative points:

- The proportion of leaking points
- $t$-test showing difference with $\infty$ of measurements

![Graph showing density of informative points]

$t$-test

$\rightarrow$ Density = 0.5
Traces Parameter: Density

**Density** of informative points:

- The proportion of leaking points
- $t$-test showing difference with $\infty$ of measurements

![Diagram showing density of informative points]

$\rightarrow$ Density = 0.9
Traces Parameter: Density

**Density** of informative points:

- The proportion of leaking points
- $t$-test showing difference with $\infty$ of measurements

Typical settings:

- Protected software: low density, long traces
- Hardware unprotected: high density, short traces

$\rightarrow$ Density = 0.9
Multi-Tuple Leakage Detection: Features

From simulations with fixed trace length:
Multi-Tuple Leakage Detection: Features

From simulations with fixed trace length:
Multi-Tuple Leakage Detection: Features

From simulations with fixed trace length:
Multi-Tuple Leakage Detection: Features

From simulations with fixed trace length:

- Both methods suffer from a low density
Multi-Tuple Leakage Detection: Features

From simulations with fixed trace length:
- Both methods suffer from a low density
- Multi-Tuple more than the TVLA
Multi-Tuple Leakage Detection: Features

From simulations with fixed trace length:

- Both methods suffer from a low density
- Multi-Tuple more than the TVLA

Reduced data complexity with higher density
Multi-Tuple Leakage Detection: Parameters

From simulations with fixed density:

\[ \log(\text{Trace length}) \]
Multi-Tuple Leakage Detection: Parameters

From simulations with fixed density:

<table>
<thead>
<tr>
<th>Trace length</th>
<th># of measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>x4</td>
</tr>
<tr>
<td>500</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>2500</td>
<td></td>
</tr>
<tr>
<td>3000</td>
<td></td>
</tr>
<tr>
<td>3500</td>
<td></td>
</tr>
</tbody>
</table>

Both methods take advantage of longer traces. Multi-Tuple gains more than the TVLA. Reduced data complexity with the number of time samples. The jointly processed trace size is limited for Hotelling's test because of covariance matrix inversion ($\sim 2000$). Possibility to run multiple Hotelling's tests in parallel.
Multi-Tuple Leakage Detection: Parameters

From simulations with fixed density:

- Both methods take advantage of longer traces.
- Multi-Tuple gains more than the TVLA.
- Reduced data complexity with the number of time samples.
- Possibility to run multiple Hotelling’s tests in parallel.
Multi-Tuple Leakage Detection: Parameters

From simulations with fixed density:
- Both methods take advantage of longer traces

![Graph showing comparison between TVLA and Multi-Tuple leakage detection methods.](image-url)
Multi-Tuple Leakage Detection: Parameters

From simulations with fixed density:

- Both methods take advantage of longer traces
- Multi-Tuple gains more than the TVLA
Multi-Tuple Leakage Detection: Parameters

From simulations with fixed density:
- Both methods take advantage of longer traces
- Multi-Tuple gains more than the TVLA

- Reduced data complexity with the number of time samples
- The jointly processed trace size is limited for Hotelling’s test because of covariance matrix inversion (∼2000):
  - Possibility to run multiple Hotelling’s tests in parallel
Two extreme settings:

White Box: everything is known about the design

Black Box: nothing is known about the design

How to perform Leakage Detection in these settings?
Practical Evaluation Scenarios

Two extreme settings:
- White Box: everything is known about the design
Practical Evaluation Scenarios

Two extreme settings:

- White Box: everything is known about the design
- Black Box: nothing is known about the design
Practical Evaluation Scenarios

Two extreme settings:
- White Box: everything is known about the design
- Black Box: nothing is known about the design

How to perform Leakage Detection in these settings?
Practical Evaluation Scenarios: White Box

In White Box:

- Prior information about leaking points
- Can reduce traces
- Can invert the covariance matrix (Hotelling’s $T^2$-test)
- High density

As a result:

- Smaller measurement period
- Easy interpretation of the confidence (no ⊥⊥ assumption)
Practical Evaluation Scenarios: White Box

In White Box:
- Prior information about leaking points

\[ P_1 \text{ or } P_2 \rightarrow C_1 \text{ or } C_2 \]
Practical Evaluation Scenarios: White Box

In White Box:
- Prior information about leaking points
- Can reduce traces

As a result:
- Smaller measurement period
- Easy interpretation of the confidence (no assumption)
Practical Evaluation Scenarios: White Box

In White Box:
- Prior information about leaking points
  - Can reduce traces
  - Can invert the covariance matrix (Hotelling’s $T^2$-test)
  - High density

As a result:
- Smaller measurement period
- Easy interpretation of the confidence (no $\perp$ assumption)
Practical Evaluation Scenarios: White Box

In White Box:
- Prior information about leaking points
  - Can reduce traces
    - Can invert the covariance matrix (Hotelling’s $T^2$-test)
    - High density

As a result:
- Smaller measurement period
- Easy interpretation of the confidence (no $\perp$ assumption)
Practical Evaluation Scenarios: Black Box

In Black Box:

- No prior information about leaking points
- Can't reduce traces
- Can't always invert the covariance matrix
- Fixed density

As a result:

- Possibly larger measurement period
- Independent assumption needed
- Heuristic required for confidence level interpretation:
  - TVLA: too conservative
  - $D$-test: too optimistic
Practical Evaluation Scenarios: Black Box

In Black Box:
- No prior information about leaking points

- Can’t reduce traces
- Can’t always invert the covariance matrix
- Fixed density

As a result:
- Possibly larger measurement period
- Independent assumption needed
- Heuristic required for confidence level interpretation:
  - TVLA: too conservative
  - D-test: too optimistic
Practical Evaluation Scenarios: Black Box

In Black Box:
- No prior information about leaking points
  - Can’t reduce traces

As a result:
- Possibly larger measurement period
- Independent assumption needed
- Heuristic required for confidence level interpretation:
  - TVLA: too conservative
  - D-test: too optimistic
Practical Evaluation Scenarios: Black Box

In Black Box:
- No prior information about leaking points
  - Can't reduce traces
  - Can't always invert the covariance matrix
Practical Evaluation Scenarios: Black Box

In Black Box:
- No prior information about leaking points
  - Can’t reduce traces
    - Can’t always invert the covariance matrix
    - Fixed density
- Possibly larger measurement period
- Independent assumption needed
- Heuristic required for confidence level interpretation:
  - TVLA: too conservative
  - $D$-test: too optimistic
Practical Evaluation Scenarios: Black Box

In Black Box:
- No prior information about leaking points
  - Can't reduce traces
    - Can't always invert the covariance matrix
    - Fixed density

As a result:
- Possibly larger measurement period
Practical Evaluation Scenarios: Black Box

In Black Box:
- No prior information about leaking points
  - Can’t reduce traces
    - Can’t always invert the covariance matrix
    - Fixed density

As a result:
- Possibly larger measurement period
- Independent assumption needed
Practical Evaluation Scenarios: Black Box

In Black Box:
- No prior information about leaking points
  - Can't reduce traces
    - Can't always invert the covariance matrix
    - Fixed density

As a result:
- Possibly larger measurement period
- Independent assumption needed
  - Heuristic required for confidence level interpretation:
    - TVLA: too conservative
    - D-test: too optimistic
Conclusion

Physical signals are not likely to be independent across time
Conclusion

Physical signals are not likely to be independent across time

1. If applicable, Hotelling’s $T^2$-test provides:
Conclusion

Physical signals are not likely to be independent across time

1. If applicable, Hotelling’s $T^2$-test provides:
   - Straight forward interpretation of the confidence level
   - And sometimes reduction the measurement period
   - Loose intuition about the POIs
Conclusion

Physical signals are not likely to be independent across time

1. If applicable, Hotelling’s $T^2$-test provides:
   ▶ Straight forward interpretation of the confidence level
   ▶ And sometimes reduction the measurement period
   ▶ Loose intuition about the POIs

2. If not, must rely on heuristics:
Conclusion

Physical signals are not likely to be independent across time

1. If applicable, Hotelling’s $T^2$-test provides:
   - Straight forward interpretation of the confidence level
   - And sometimes reduction the measurement period
   - Loose intuition about the POIs

2. If not, must rely on heuristics:
   - TVLA: too conservative
   - $D$-test: too optimistic
Conclusion

Physical signals are not likely to be independent across time

1. If applicable, Hotelling’s $T^2$-test provides:
   - Straight forward interpretation of the confidence level
   - And sometimes reduction the measurement period
   - Loose intuition about the POIs

2. If not, must rely on heuristics:
   - TVLA: too conservative
   - $D$-test: too optimistic

Evaluation Hardness
Conclusion

Physical signals are not likely to be independent across time

1. If applicable, Hotelling’s $T^2$-test provides:
   - Straight forward interpretation of the confidence level
   - And sometimes reduction the measurement period
   - Loose intuition about the POIs

2. If not, must rely on heuristics:
   - TVLA: too conservative
   - $D$-test: too optimistic

Thanks!

Evaluation Hardness

github.com/obronchain/multituple_leakage_detection