Vorsicht Hintertür!

Seitenkanäle als versteckte Systemschwachstellen

Markus Kasper

Lehrstuhl für Embedded Security
Ruhr-Universität Bochum

Labor Vortrag 2010

Outline

1. Einführung
2. Einfache Angriffe: SPA und DPA
3. Stromverbrauch und Modelle
4. Ein wenig Statistik
5. CPA - Correlation Power Analysis
6. Summary

Einführung

Wo sind wir / Worum gehts heute Abend?

Wir reden über:

Wo sind wir / Worum gehts heute Abend?

Wir reden über:

- IT-Sicherheit

Wo sind wir / Worum gehts heute Abend?

Wir reden über:

- IT-Sicherheit
- Eingebettete Systeme (z.B. Mikrocontroller)

Wo sind wir / Worum gehts heute Abend?

Wir reden über:

- IT-Sicherheit
- Eingebettete Systeme (z.B. Mikrocontroller)
- Angriffe, welche nicht die kryptografischen Algorithmen angreifen, sondern die Tatsache, dass sie in echten Systemen implementiert sind

Wo sind wir / Worum gehts heute Abend?

Wir reden über:

- IT-Sicherheit
- Eingebettete Systeme (z.B. Mikrocontroller)
- Angriffe, welche nicht die kryptografischen Algorithmen angreifen, sondern die Tatsache, dass sie in echten Systemen implementiert sind
- Können auch hochsichere Algorithmen aushebeln, selbst wenn sie mathematisch “beweisbar sicher” sind.

Wo sind wir / Worum gehts heute Abend?

Wir reden über:

- IT-Sicherheit
- Eingebettete Systeme (z.B. Mikrocontroller)
- Angriffe, welche nicht die kryptografischen Algorithmen angreifen, sondern die Tatsache, dass sie in echten Systemen implementiert sind
- Können auch hochsichere Algorithmen aushebeln, selbst wenn sie mathematisch “beweisbar sicher” sind.
- Kein Rocket-Science: nur Signal-Processing

Wo sind wir / Worum gehts heute Abend?

Wir reden über:

- IT-Sicherheit
- Eingebettete Systeme (z.B. Mikrocontroller)
- Angriffe, welche nicht die kryptografischen Algorithmen angreifen, sondern die Tatsache, dass sie in echten Systemen implementiert sind
- Können auch hochsichere Algorithmen aushebeln, selbst wenn sie mathematisch “beweisbar sicher” sind.
- Kein Rocket-Science: nur Signal-Processing
- Heute: Nur passive Angriffe (Seitenkanalanalysen)

Hacker Paragraph

ACHTUNG:
Viele Inhalte des Abends erläutern Sachverhalte, deren missbräuchliche Anwendung (sprich: auf fremde Systeme) in Deutschland strafbar ist oder im Graubereich liegt.

Hacker Paragraph

ACHTUNG:
Viele Inhalte des Abends erläutern Sachverhalte, deren missbräuchliche Anwendung (sprich: auf fremde Systeme) in Deutschland strafbar ist oder im Graubereich liegt.

Handelt bitte entsprechend Verantwortungsbewusst!

Bedrohungspotential

- Die Angriffe sind seit etwa 1999 bekannt

Bedrohungspotential

- Die Angriffe sind seit etwa 1999 bekannt
- Es gibt Abwehrmaßnahmen

Bedrohungspotential

- Die Angriffe sind seit etwa 1999 bekannt
- Es gibt Abwehrmaßnahmen
- Diese taugen meist nichts

Bedrohungspotential

- Die Angriffe sind seit etwa 1999 bekannt
- Es gibt Abwehrmaßnahmen
- Diese taugen meist nichts
- Lösung: Ein Seitenkanalangriff darf keinen Vorteil bringen

Bedrohungspotential

Beispiel KeeLoq:

- Es gibt Handsender und Empfänger

Bedrohungspotential

Beispiel KeeLoq:

- Es gibt Handsender und Empfänger
- Jeder Handsender hat einen eigenen geheimen Schlüssel

Bedrohungspotential

Beispiel KeeLoq:

- Es gibt Handsender und Empfänger
- Jeder Handsender hat einen eigenen geheimen Schlüssel
- Die Handsender-Schlüssel können am Empfänger aus der Seriennummer und einem globalen Schlüssel abgeleitet werden (= Verschlüsselung)

Bedrohungspotential

Beispiel KeeLoq:

- Es gibt Handsender und Empfänger
- Jeder Handsender hat einen eigenen geheimen Schlüssel
- Die Handsender-Schlüssel können am Empfänger aus der Seriennummer und einem globalen Schlüssel abgeleitet werden (= Verschlüsselung)
- Wer den globalen Schlüssel hat kann nach Abhören einer einzigen Nachricht **alle** Tore des Herstellers öffnen

Bedrohungspotential

Beispiel KeeLoq:

- Es gibt Handsender und Empfänger
- Jeder Handsender hat einen eigenen geheimen Schlüssel
- Die Handsender-Schlüssel können am Empfänger aus der Seriennummer und einem globalen Schlüssel abgeleitet werden (= Verschlüsselung)
- Wer den globalen Schlüssel hat kann nach Abhören einer einzigen Nachricht **alle** Tore des Herstellers öffnen
- Klingt schrecklich, aber in Wirklichkeit benutzen viele solcher Systeme gar keine Verschlüsselung und senden immer den gleichen Code

Bedrohungspotential

Beispiel KeeLoq:

- Es gibt Handsender und Empfänger
- Jeder Handsender hat einen eigenen geheimen Schlüssel
- Die Handsender-Schlüssel können am Empfänger aus der Seriennummer und einem globalen Schlüssel abgeleitet werden (= Verschlüsselung)
- Wer den globalen Schlüssel hat kann nach Abhören einer einzigen Nachricht **alle** Tore des Herstellers öffnen
- Klingt schrecklich, aber in Wirklichkeit benutzen viele solcher Systeme gar keine Verschlüsselung und senden immer den gleichen Code
- Und trotzdem lesen wir nicht täglich von entsperrten Garagen!

Bedrohungspotential

Beispiel KeeLoq:

- Es gibt Handsender und Empfänger
- Jeder Handsender hat einen eigenen geheimen Schlüssel
- Die Handsender-Schlüssel können am Empfänger aus der Seriennummer und einem globalen Schlüssel abgeleitet werden (= Verschlüsselung)
- Wer den globalen Schlüssel hat kann nach Abhören einer einzigen Nachricht **alle** Tore des Herstellers öffnen
- Klingt schrecklich, aber in Wirklichkeit benutzen viele solcher Systeme gar keine Verschlüsselung und senden immer den gleichen Code
- Und trotzdem lesen wir nicht täglich von entsperrten Garagen!
- Wozu Seitenkanalangriffe, wenn es das Brecheisen auch tut?

Was ist ein Seitenkanal?

- Ein Seitenkanal ist eine messbare physikalische Größe, die es erlaubt Informationen über ein System zu extrahieren.

Was ist ein Seitenkanal?

- Ein Seitenkanal ist eine messbare physikalische Größe, die es erlaubt Informationen über ein System zu extrahieren.
- Es wird ausgenutzt, dass **reale** Systeme unvermeidbar immer ihre eigenen physikalische Eigenschaften und Verhaltensweisen besitzen.

Was ist ein Seitenkanal?

- Ein Seitenkanal ist eine messbare physikalische Größe, die es erlaubt Informationen über ein System zu extrahieren.
- Es wird ausgenutzt, dass **reale** Systeme unvermeidbar immer ihre eigenen physikalische Eigenschaften und Verhaltensweisen besitzen.
- Bei uns: Informationen = Krypto-Schlüssel

Was ist ein Seitenkanal?

- Ein Seitenkanal ist eine messbare physikalische Größe, die es erlaubt Informationen über ein System zu extrahieren.
- Es wird ausgenutzt, dass **reale** Systeme unvermeidbar immer ihre eigenen physikalische Eigenschaften und Verhaltensweisen besitzen.
- Bei uns: Informationen = Krypto-Schlüssel
- Seitenkanäle werden auch in anderen Industrien eingesetzt: z.B. als Ersatz für Sensoren (z.B. Kfz Elektronik)

Was ist ein Seitenkanal?

- Ein Seitenkanal ist eine messbare physikalische Größe, die es erlaubt Informationen über ein System zu extrahieren.
- Es wird ausgenutzt, dass **reale** Systeme unvermeidbar immer ihre eigenen physikalische Eigenschaften und Verhaltensweisen besitzen.
- Bei uns: Informationen = Krypto-Schlüssel
- Seitenkanäle werden auch in anderen Industrien eingesetzt: z.B. als Ersatz für Sensoren (z.B. Kfz Elektronik)

Seitenkanalangriffe greifen die **Implementierung** an, nicht den Algorithmus!

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

- Ampelbutton

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

- Ampelbutton
- Licht

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

- Ampelbutton
- Licht
- Tür

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

- Ampelbutton
- Licht
- Tür
- Pizzabestellungen

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

- Ampelbutton
- Licht
- Tür
- Pizzabestellungen
- Autos

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

- Ampelbutton
- Licht
- Tür
- Pizzabestellungen
- Autos
- Lautstärke

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

- Ampelbutton
- Licht
- Tür
- Pizzabestellungen
- Autos
- Lautstärke
- Anzahl der WLAN Devices

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

- Ampelbutton
- Licht
- Tür
- Pizzabestellungen
- Autos
- Lautstärke
- Anzahl der WLAN Devices
- Stromverbrauch

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

- Ampelbutton
- Licht
- Tür
- Pizzabestellungen
- Autos
- Lautstärke
- Anzahl der WLAN Devices
- Stromverbrauch
- Temperatur

Beispiele für Seitenkanäle

Frage: Ist Jmd. im Labor?
Noch besser: Wie viele Menschen sind im Labor?

- Ampelbutton
- Licht
- Tür
- Pizzabestellungen
- Autos
- Lautstärke
- Anzahl der WLAN Devices
- Stromverbrauch
- Temperatur
- Geruch

Typische Seitenkanäle

Für praktische Angriffe auf elektronische Geräte verwendet man in der Regel

- Zeit

Typische Seitenkanäle

Für praktische Angriffe auf elektronische Geräte verwendet man in der Regel

- Zeit
- EM Abstrahlung

Typische Seitenkanäle

Für praktische Angriffe auf elektronische Geräte verwendet man in der Regel

- Zeit
- EM Abstrahlung
- Stromverbrauch

Grundsätzliche Funktionsweise der Angriffe

1. Divide-and-Conquer Strategie!

Grundsätzliche Funktionsweise der Angriffe

1. Divide-and-Conquer Strategie!
2. Die Größe (z.B. Stromverbrauch) wird aufgezeichnet während der Alg. ausgeführt wird
   - Unterschiedliche (zufällige) Eingabewerte
   - Speichern: Kurve & Eingabewert
   - Manchmal: Ausgabewerte statt Eingabewerte

Grundsätzliche Funktionsweise der Angriffe

1. Divide-and-Conquer Strategie!
2. Die Größe (z.B. Stromverbrauch) wird aufgezeichnet während der Alg. ausgeführt wird
   - Unterschiedliche (zufällige) Eingabewerte
   - Speichern: Kurve & Eingabewert
   - Manchmal: Ausgabewerte statt Eingabewerte
3. Evtl. etwas Preprocessing der Messdaten
   - Kompression
   - Filtern
   - Alignment

Grundsätzliche Funktionsweise der Angriffe

4. Zwischenwerte $= f(Input, SubKey)$ unter Annahme eines *SubKey*s vorhersagen
   - Algorithmus muss bekannt sein
   - *SubKey* steht für einen kleinen Teil des Schlüssels bzw. Rundenschlüssels.
   - Klein bedeutet in der Praxis weniger als 10 Bit
   - Zwischenwert nichtlinear abhängig von *SubKey*

Grundsätzliche Funktionsweise der Angriffe

4. Zwischenwerte $= f(Input, SubKey)$ unter Annahme eines *SubKey*s vorhersagen
   - Algorithmus muss bekannt sein
   - *SubKey* steht für einen kleinen Teil des Schlüssels bzw. Rundenschlüssels.
   - Klein bedeutet in der Praxis weniger als 10 Bit
   - Zwischenwert nichtlinear abhängig von *SubKey*
5. Die Vorhersagen der Zwischenwerte mit der gemessenen Größe in Zusammenhang bringen

Grundsätzliche Funktionsweise der Angriffe

4. Zwischenwerte $= f(Input, SubKey)$ unter Annahme eines *SubKey*s vorhersagen
   - Algorithmus muss bekannt sein
   - *SubKey* steht für einen kleinen Teil des Schlüssels bzw. Rundenschlüssels.
   - Klein bedeutet in der Praxis weniger als 10 Bit
   - Zwischenwert nichtlinear abhängig von *SubKey*
5. Die Vorhersagen der Zwischenwerte mit der gemessenen Größe in Zusammenhang bringen
6. Vorhersagen mit Messungen vergleichen

Grundsätzliche Funktionsweise der Angriffe

4. Zwischenwerte $= f(Input, SubKey)$ unter Annahme eines *SubKey*s vorhersagen
   - Algorithmus muss bekannt sein
   - *SubKey* steht für einen kleinen Teil des Schlüssels bzw. Rundenschlüssels.
   - Klein bedeutet in der Praxis weniger als 10 Bit
   - Zwischenwert nichtlinear abhängig von *SubKey*
5. Die Vorhersagen der Zwischenwerte mit der gemessenen Größe in Zusammenhang bringen
6. Vorhersagen mit Messungen vergleichen
7. Beim richtigen SubKey “passt” es am besten

Einfache Angriffe: SPA und DPA

Messaufbau

Am häufigsten wird der Stromverbrauch zur Seitenkanalanalyse verwendet:

- Ein PC löst Verschlüsselungen aus und zeichnet Ein-/Ausgabe auf.

Messaufbau

Am häufigsten wird der Stromverbrauch zur Seitenkanalanalyse verwendet:

- Ein PC löst Verschlüsselungen aus und zeichnet Ein-/Ausgabe auf.
- Ein Oszilloskop zeichnet den Stromverbrauch auf

Messaufbau

Am häufigsten wird der Stromverbrauch zur Seitenkanalanalyse verwendet:

- Ein PC löst Verschlüsselungen aus und zeichnet Ein-/Ausgabe auf.
- Ein Oszilloskop zeichnet den Stromverbrauch auf
- Gemessen wird über einen kleinen Widerstand zwischen GND des Devices und GND der Spannungsversorgung.

Messaufbau

Am häufigsten wird der Stromverbrauch zur Seitenkanalanalyse verwendet:

- Ein PC löst Verschlüsselungen aus und zeichnet Ein-/Ausgabe auf.
- Ein Oszilloskop zeichnet den Stromverbrauch auf
- Gemessen wird über einen kleinen Widerstand zwischen GND des Devices und GND der Spannungsversorgung.
- Die Messung wird zwischen 50 und 1 Mio. mal durchgeführt.

SPA - Simple Power Analysis

- Eine Simple Power Analysis verwendet nur einen einzigen oder wenige gemittelte Messkurven

SPA - Simple Power Analysis

- Eine Simple Power Analysis verwendet nur einen einzigen oder wenige gemittelte Messkurven
- Die gewünschte Information lässt sich direkt ablesen

SPA - Simple Power Analysis

- Eine Simple Power Analysis verwendet nur einen einzigen oder wenige gemittelte Messkurven
- Die gewünschte Information lässt sich direkt ablesen
- Sehr einfach

SPA - Simple Power Analysis

- Eine Simple Power Analysis verwendet nur einen einzigen oder wenige gemittelte Messkurven
- Die gewünschte Information lässt sich direkt ablesen
- Sehr einfach
- Beispiel: Square-and-Multiply

SPA - Simple Power Analysis

Beispiel: Stromkurve (MATLAB)

DPA - Differential Power Analysis

Bisher:

- Viele Informationen, aber Schlüssel höchstens bei SAM

DPA - Differential Power Analysis

Bisher:

- Viele Informationen, aber Schlüssel höchstens bei SAM
- Nur einzelne Traces

DPA - Differential Power Analysis

Bisher:

- Viele Informationen, aber Schlüssel höchstens bei SAM
- Nur einzelne Traces
- Eingabe und Zwischenwerte egal

DPA - Differential Power Analysis

Bisher:

- Viele Informationen, aber Schlüssel höchstens bei SAM
- Nur einzelne Traces
- Eingabe und Zwischenwerte egal

Jetzt:

- Verwendung vieler Kurven

DPA - Differential Power Analysis

Bisher:

- Viele Informationen, aber Schlüssel höchstens bei SAM
- Nur einzelne Traces
- Eingabe und Zwischenwerte egal

Jetzt:

- Verwendung vieler Kurven
- Verwendung einzelner Messpunkte

DPA - Differential Power Analysis

Bisher:

- Viele Informationen, aber Schlüssel höchstens bei SAM
- Nur einzelne Traces
- Eingabe und Zwischenwerte egal

Jetzt:

- Verwendung vieler Kurven
- Verwendung einzelner Messpunkte
- Verwendung der Vorhersagen

DPA - Differential Power Analysis

Idee:

- Unterschiedliche Zwischenwerte resultieren in einem unterschiedlichem Stromverbrauch

DPA - Differential Power Analysis

Idee:

- Unterschiedliche Zwischenwerte resultieren in einem unterschiedlichem Stromverbrauch
- Mitteln über viele Kurven: Kurve, die repräsentativ für Gemeinsamkeiten der gemittelten Kurven ist

DPA - Differential Power Analysis

Idee:

- Unterschiedliche Zwischenwerte resultieren in einem unterschiedlichem Stromverbrauch
- Mitteln über viele Kurven: Kurve, die repräsentativ für Gemeinsamkeiten der gemittelten Kurven ist
- Kurven in 2 Fälle sortieren:
  z.B. Bit gesetzt vs. Bit nicht gesetzt
  oder alle Bits gesetzt vs. kein Bit gesetzt.

DPA - Differential Power Analysis

Idee:

- Unterschiedliche Zwischenwerte resultieren in einem unterschiedlichem Stromverbrauch
- Mitteln über viele Kurven: Kurve, die repräsentativ für Gemeinsamkeiten der gemittelten Kurven ist
- Kurven in 2 Fälle sortieren:
  z.B. Bit gesetzt vs. Bit nicht gesetzt
  oder alle Bits gesetzt vs. kein Bit gesetzt.
- Für jede Gruppe die Mittelwertkurve bilden

DPA - Differential Power Analysis

Idee:

- Unterschiedliche Zwischenwerte resultieren in einem unterschiedlichem Stromverbrauch
- Mitteln über viele Kurven: Kurve, die repräsentativ für Gemeinsamkeiten der gemittelten Kurven ist
- Kurven in 2 Fälle sortieren:
  z.B. Bit gesetzt vs. Bit nicht gesetzt
  oder alle Bits gesetzt vs. kein Bit gesetzt.
- Für jede Gruppe die Mittelwertkurve bilden
- Die beiden Kurven von einander abziehen

DPA - Differential Power Analysis

Ergebnis:

- War die Sortierung sinnvoll, unterscheiden sich die Mittelwerte der Kurven

DPA - Differential Power Analysis

Ergebnis:

- War die Sortierung sinnvoll, unterscheiden sich die Mittelwerte der Kurven
- ... und zwar an allen Stellen, an denen sich die Aufteiung bemerkbar macht

DPA - Differential Power Analysis

Ergebnis:

- War die Sortierung sinnvoll, unterscheiden sich die Mittelwerte der Kurven
- ... und zwar an allen Stellen, an denen sich die Aufteiung bemerkbar macht
- Der Zeitpunkt, an dem der Zwischenwert verwendet wurde ist Teil des Ergebnisses!

DPA - Differential Power Analysis

Ergebnis:

- War die Sortierung sinnvoll, unterscheiden sich die Mittelwerte der Kurven
- ... und zwar an allen Stellen, an denen sich die Aufteiung bemerkbar macht
- Der Zeitpunkt, an dem der Zwischenwert verwendet wurde ist Teil des Ergebnisses!
- Benötigt werden nur Ein- oder Ausgabe und Algorithmus

DPA - Differential Power Analysis

Ergebnis:

- War die Sortierung sinnvoll, unterscheiden sich die Mittelwerte der Kurven
- ... und zwar an allen Stellen, an denen sich die Aufteiung bemerkbar macht
- Der Zeitpunkt, an dem der Zwischenwert verwendet wurde ist Teil des Ergebnisses!
- Benötigt werden nur Ein- oder Ausgabe und Algorithmus
- Sehr einfach, sehr effizient, sehr leicht verständlich

DPA - Differential Power Analysis

Ergebnis:

- War die Sortierung sinnvoll, unterscheiden sich die Mittelwerte der Kurven
- ... und zwar an allen Stellen, an denen sich die Aufteiung bemerkbar macht
- Der Zeitpunkt, an dem der Zwischenwert verwendet wurde ist Teil des Ergebnisses!
- Benötigt werden nur Ein- oder Ausgabe und Algorithmus
- Sehr einfach, sehr effizient, sehr leicht verständlich
- Sehr weit weg von einer cleveren Lösung

Ein paar Hinweise zur Praxis

Wo greift man an?

Ein paar Hinweise zur Praxis

Wo greift man an?

Der Zwischenwert sollte nichtlinear vom Schlüssel abhängen.
(Begründung kommt beim Power Model!)

Ein paar Hinweise zur Praxis

Wo greift man an?

Der Zwischenwert sollte nichtlinear vom Schlüssel abhängen.
(Begründung kommt beim Power Model!)

Wie lang ist eine Trace?

Ein paar Hinweise zur Praxis

Wo greift man an?

Der Zwischenwert sollte nichtlinear vom Schlüssel abhängen.
(Begründung kommt beim Power Model!)

Wie lang ist eine Trace?

Hängt von der Sample Rate ab und davon wie genau man die korrekte Position des Angriffspunkts abschätzen kann.
Von 500 bis 8 Mio. Messwerten ist alles drin

Ein paar Hinweise zur Praxis

Wo greift man an?

Der Zwischenwert sollte nichtlinear vom Schlüssel abhängen.
(Begründung kommt beim Power Model!)

Wie lang ist eine Trace?

Hängt von der Sample Rate ab und davon wie genau man die korrekte Position des Angriffspunkts abschätzen kann.
Von 500 bis 8 Mio. Messwerten ist alles drin

Wie viele Messungen werden benötigt?

Ein paar Hinweise zur Praxis

Wo greift man an?
Der Zwischenwert sollte nichtlinear vom Schlüssel abhängen.
(Begründung kommt beim Power Model!)

Wie lang ist eine Trace?
Hängt von der Sample Rate ab und davon wie genau man die korrekte Position des Angriffspunkts abschätzen kann.
Von 500 bis 8 Mio. Messwerten ist alles drin

Wie viele Messungen werden benötigt?
Hängt einzig und allein vom Device und der Messqualität ab. Gute Angriffe weniger als 100 Kurven, Mikrocontroller etwa 1000, schlechte Angriffe 10K-20K

Ein paar Hinweise zur Praxis

Wo greift man an?
Der Zwischenwert sollte nichtlinear vom Schlüssel abhängen.
(Begründung kommt beim Power Model!)

Wie lang ist eine Trace?
Hängt von der Sample Rate ab und davon wie genau man die korrekte Position des Angriffspunkts abschätzen kann.
Von 500 bis 8 Mio. Messwerten ist alles drin

Wie viele Messungen werden benötigt?
Hängt einzig und allein vom Device und der Messqualität ab. Gute Angriffe weniger als 100 Kurven, Mikrocontroller etwa 1000, schlechte Angriffe 10K-20K

Das sind eine Menge Daten!

Ein paar Hinweise zur Praxis

Wo greift man an?
Der Zwischenwert sollte nichtlinear vom Schlüssel abhängen. (Begründung kommt beim Power Model!)

Wie lang ist eine Trace?
Hängt von der Sample Rate ab und davon wie genau man die korrekte Position des Angriffspunkts abschätzen kann.
Von 500 bis 8 Mio. Messwerten ist alles drin

Wie viele Messungen werden benötigt?
Hängt einzig und allein vom Device und der Messqualität ab. Gute Angriffe weniger als 100 Kurven, Mikrocontroller etwa 1000, schlechte Angriffe 10K-20K

Das sind eine Menge Daten!
Ja! Deshalb: Nach Möglichkeit Komprimieren (Peak Extraction oder Integrieren)

Ein paar Hinweise zur Praxis

Wo greift man an?
Der Zwischenwert sollte nichtlinear vom Schlüssel abhängen. (Begründung kommt beim Power Model!)

Wie lang ist eine Trace?
Hängt von der Sample Rate ab und davon wie genau man die korrekte Position des Angriffspunkts abschätzen kann.
Von 500 bis 8 Mio. Messwerten ist alles drin

Wie viele Messungen werden benötigt?
Hängt einzig und allein vom Device und der Messqualität ab. Gute Angriffe weniger als 100 Kurven, Mikrocontroller etwa 1000, schlechte Angriffe 10K-20K

Das sind eine Menge Daten!
Ja! Deshalb: Nach Möglichkeit Komprimieren (Peak Extraction oder Integrieren)

Was noch?

Ein paar Hinweise zur Praxis

Wo greift man an?
Der Zwischenwert sollte nichtlinear vom Schlüssel abhängen. (Begründung kommt beim Power Model!)

Wie lang ist eine Trace?
Hängt von der Sample Rate ab und davon wie genau man die korrekte Position des Angriffspunkts abschätzen kann.
Von 500 bis 8 Mio. Messwerten ist alles drin

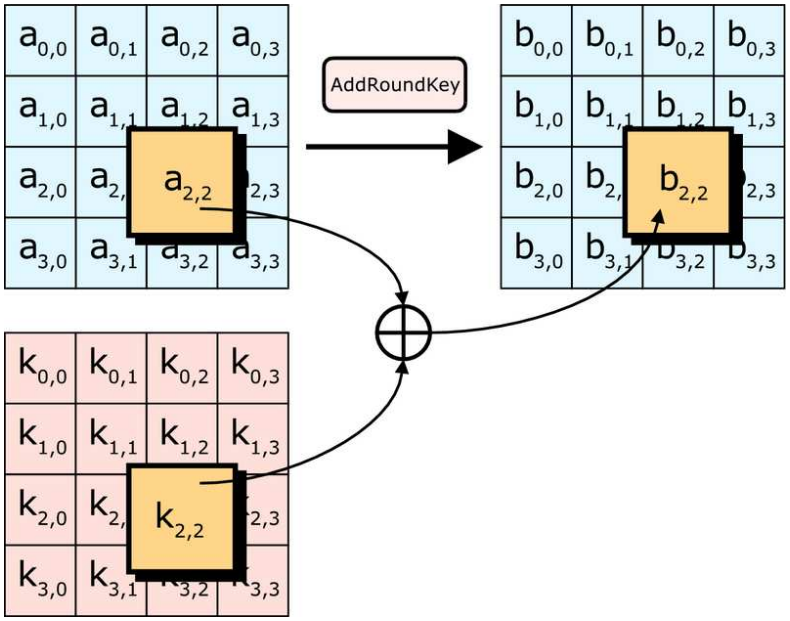
Wie viele Messungen werden benötigt?
Hängt einzig und allein vom Device und der Messqualität ab. Gute Angriffe weniger als 100 Kurven, Mikrocontroller etwa 1000, schlechte Angriffe 10K-20K

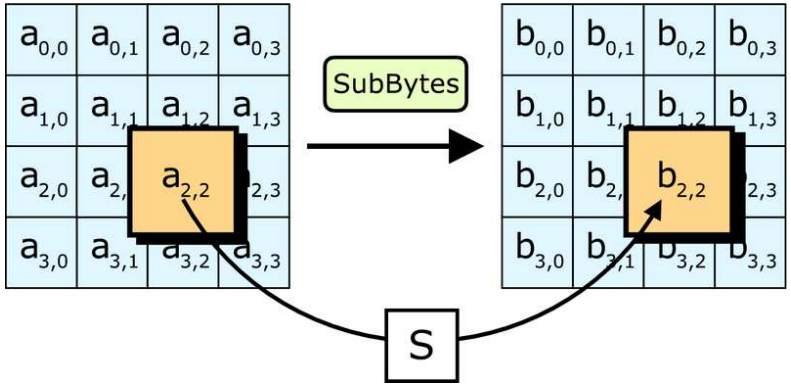
Das sind eine Menge Daten!
Ja! Deshalb: Nach Möglichkeit Komprimieren (Peak Extraction oder Integrieren)

Was noch?
Alignment ist Mega wichtig! Sonst werden die falschen Messwerte verwendet!

DPA - Differential Power Analysis

Beispiel: Stromkurve AES (MATLAB)





Stromverbrauch und Modelle

Was bewirkt den Stromverbrauch?

- Schaltungen verwenden CMOS Logik

Was bewirkt den Stromverbrauch?

- Schaltungen verwenden CMOS Logik
- CMOS Logik besteht aus einem pull-up und einem pull-down Netzwerk aus Transistoren, wovon immer nur eines zur Versorgungsspannung oder zu GND verbindet.

Was bewirkt den Stromverbrauch?

- Schaltungen verwenden CMOS Logik
- CMOS Logik besteht aus einem pull-up und einem pull-down Netzwerk aus Transistoren, wovon immer nur eines zur Versorgungsspannung oder zu GND verbindet.
- CMOS Inverter (Tafel)

Was bewirkt den Stromverbrauch?

- Schaltungen verwenden CMOS Logik
- CMOS Logik besteht aus einem pull-up und einem pull-down Netzwerk aus Transistoren, wovon immer nur eines zur Versorgungsspannung oder zu GND verbindet.
- CMOS Inverter (Tafel)

1. Kurzschlussströme (dynamisch)

Was bewirkt den Stromverbrauch?

- Schaltungen verwenden CMOS Logik
- CMOS Logik besteht aus einem pull-up und einem pull-down Netzwerk aus Transistoren, wovon immer nur eines zur Versorgungsspannung oder zu GND verbindet.
- CMOS Inverter (Tafel)

1. Kurzschlussströme (dynamisch)
2. Umladung der Kapazitäten (dynamisch)

Was bewirkt den Stromverbrauch?

- Schaltungen verwenden CMOS Logik
- CMOS Logik besteht aus einem pull-up und einem pull-down Netzwerk aus Transistoren, wovon immer nur eines zur Versorgungsspannung oder zu GND verbindet.
- CMOS Inverter (Tafel)

1. Kurzschlussströme (dynamisch)
2. Umladung der Kapazitäten (dynamisch)
3. Statische Leckströme (z.B. durchs Gate)

Was bewirkt den Stromverbrauch?

Zusammenfassung:

- Wichtig ist der dynamische Anteil

Was bewirkt den Stromverbrauch?

Zusammenfassung:

- Wichtig ist der dynamische Anteil
- Nur $0 \rightarrow 1$ Übergänge führen zu einem Ladestrom für die Kapazität

Was bewirkt den Stromverbrauch?

Zusammenfassung:

- Wichtig ist der dynamische Anteil
- Nur $0 \rightarrow 1$ Übergänge führen zu einem Ladestrom für die Kapazität
- Anzahl Glitches (Kurzschlussströme) summieren

Hamming Weight und Hamming Distance

Ideal:
Wir haben alle Infos um eine Simulation des Stromverbrauchs auf Transistorebene durchzuführen.

Hamming Weight und Hamming Distance

Ideal:
Wir haben alle Infos um eine Simulation des Stromverbrauchs auf Transistorebene durchzuführen.

Real:
Wir haben keine Ahnung welcher Code ausgeführt wird und wie genau die Hardware verschaltet ist.

Hamming Weight und Hamming Distance

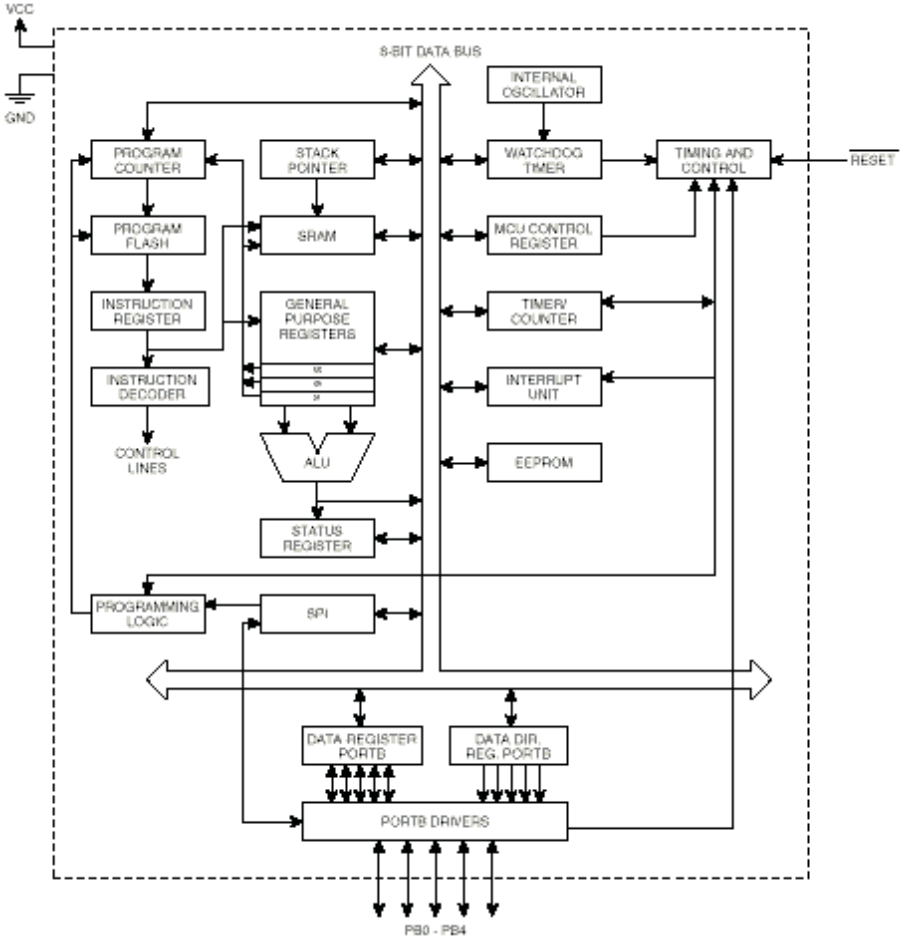
Ideal:
Wir haben alle Infos um eine Simulation des Stromverbrauchs auf Transistorebene durchzuführen.

Real:
Wir haben keine Ahnung welcher Code ausgeführt wird und wie genau die Hardware verschaltet ist.

Lösung:
Einfache Modelle!

Hamming Weight und Hamming Distance

Blockschaltbild des AT90S2323/2343



Hamming Weight und Hamming Distance

Feststellung 1:

- Der Datenbus in einem Mikrocontroller oder auch im ASIC ist riesig lang und führt zu vielen Komponenten

Hamming Weight und Hamming Distance

Feststellung 1:

- Der Datenbus in einem Mikrocontroller oder auch im ASIC ist riesig lang und führt zu vielen Komponenten
- Dann wird er auch eine große Kapazität haben!

Hamming Weight und Hamming Distance

Feststellung 1:

- Der Datenbus in einem Mikrocontroller oder auch im ASIC ist riesig lang und führt zu vielen Komponenten
- Dann wird er auch eine große Kapazität haben!
- Nur wenn eine Datenleitung auf 1 liegt, **kann** ein Ladestrom geflossen sein.

Hamming Weight und Hamming Distance

Feststellung 1:

- Der Datenbus in einem Mikrocontroller oder auch im ASIC ist riesig lang und führt zu vielen Komponenten
- Dann wird er auch eine große Kapazität haben!
- Nur wenn eine Datenleitung auf 1 liegt, **kann** ein Ladestrom geflossen sein.
- Wir zählen die Anzahl der 1 en auf dem Bus

Hamming Weight und Hamming Distance

Feststellung 1:

- Der Datenbus in einem Mikrocontroller oder auch im ASIC ist riesig lang und führt zu vielen Komponenten
- Dann wird er auch eine große Kapazität haben!
- Nur wenn eine Datenleitung auf 1 liegt, **kann** ein Ladestrom geflossen sein.
- Wir zählen die Anzahl der 1 en auf dem Bus

Hamming Weight Modell:

$$P = \sum_{i=1}^{8} u_i$$

Hamming Weight und Hamming Distance

Feststellung 2:

- Umladen führt zu Glitches

Hamming Weight und Hamming Distance

Feststellung 2:

- Umladen führt zu Glitches
- Kennt man den vorherigen Zustand kann man mit den Glitches arbeiten

Hamming Weight und Hamming Distance

Feststellung 2:

- Umladen führt zu Glitches
- Kennt man den vorherigen Zustand kann man mit den Glitches arbeiten
- Anzahl der Bitwechsel in einem Takt zählen (=XOR Vorher/Nachher)

Hamming Weight und Hamming Distance

Feststellung 2:

- Umladen führt zu Glitches
- Kennt man den vorherigen Zustand kann man mit den Glitches arbeiten
- Anzahl der Bitwechsel in einem Takt zählen (=XOR Vorher/Nachher)

Hamming Distance Modell:

$$P = \sum_{i=1}^{8} (u_i \oplus v_i)$$

Hamming Weight und Hamming Distance

Zusammenfassung:

- HW Model funktioniert gut bei Mikrocontrollern

Hamming Weight und Hamming Distance

Zusammenfassung:

- HW Model funktioniert gut bei Mikrocontrollern
- Bei ASICS HW Model oder HD Model (hängt vom Chip ab)

Hamming Weight und Hamming Distance

Zusammenfassung:

- HW Model funktioniert gut bei Mikrocontrollern
- Bei ASICS HW Model oder HD Model (hängt vom Chip ab)
- Nichtlineare Abhängigkeit bedeutet völlig unterschiedlicher Stromverbrauch bei ähnlichen Schlüsseln!

Hamming Weight und Hamming Distance

Zusammenfassung:

- HW Model funktioniert gut bei Mikrocontrollern
- Bei ASICS HW Model oder HD Model (hängt vom Chip ab)
- Nichtlineare Abhängigkeit bedeutet völlig unterschiedlicher Stromverbrauch bei ähnlichen Schlüsseln!
- “Alles Null” und “Alles Eins” sind am einfachsten unterscheidbar

Ein wenig Statistik

Ein wenig Statistik

Achtung Formeln! Bitte nicht weglaufen oder einschlafen!

Mittelwert und Varianz

Mittelwert:

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

- Enthält keine Information
- Wird entfernt
- Hilft Alignment zu testen und Strukturen zu erkennen

Mittelwert und Varianz

Varianz:

$$\mathrm{Var}(x) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

- Wie stark verändert sich der Messwert?
- Enthält Information
- Keine Varianz bedeutet keine Information

Mittelwert und Varianz

Beispiel: Interrupt-Routine entfernen

Korrelations-Koeffizient

(Pearsons Correlation Coefficient)

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

- Standard Tool für Seitenkanalangriffe

Korrelations-Koeffizient

(Pearsons Correlation Coefficient)

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

- Standard Tool für Seitenkanalangriffe
- Ein Maß für lineare Zusammenhänge

Korrelations-Koeffizient

(Pearsons Correlation Coefficient)

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

- Standard Tool für Seitenkanalangriffe
- Ein Maß für lineare Zusammenhänge
- Ignoriert den Mittelwert!

Korrelations-Koeffizient

(Pearsons Correlation Coefficient)

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

- Standard Tool für Seitenkanalangriffe
- Ein Maß für lineare Zusammenhänge
- Ignoriert den Mittelwert!
- Normalisiert zwischen -1 und 1

Korrelations-Koeffizient

(Pearsons Correlation Coefficient)

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

- Standard Tool für Seitenkanalangriffe
- Ein Maß für lineare Zusammenhänge
- Ignoriert den Mittelwert!
- Normalisiert zwischen -1 und 1
- Schreckliche Formel, kann jedoch einfach verwendet werden

Korrelations-Koeffizient

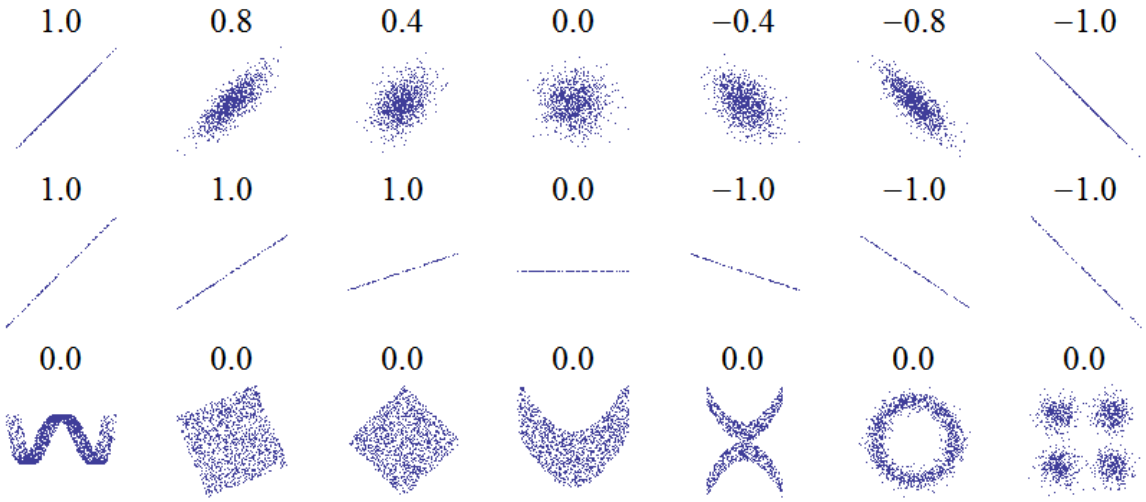
(Pearsons Correlation Coefficient)

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

- Standard Tool für Seitenkanalangriffe
- Ein Maß für lineare Zusammenhänge
- Ignoriert den Mittelwert!
- Normalisiert zwischen -1 und 1
- Schreckliche Formel, kann jedoch einfach verwendet werden
- 1 = linear, -1 = linear mit neg. Faktor, 0 = kein Zusammenhang

Korrelations-Koeffizient

Von Wikipedia geklaute Beispiele:



Hinweise zu Angriffen mit Korrelations-Koeffizient

Wie hoch ist der Korrelations-Koeffizient?

Hinweise zu Angriffen mit Korrelations-Koeffizient

Wie hoch ist der Korrelations-Koeffizient?

$$\rho = \sqrt{\frac{\mathrm{SNR}}{\mathrm{SNR}+1}} = \sqrt{\frac{1}{1+\mathrm{SNR}^{-1}}}$$

Wie viele Kurven werden benötigt?

Hinweise zu Angriffen mit Korrelations-Koeffizient

Wie hoch ist der Korrelations-Koeffizient?

$$\rho = \sqrt{\frac{\mathrm{SNR}}{\mathrm{SNR}+1}} = \sqrt{\frac{1}{1+\mathrm{SNR}^{-1}}}$$

Wie viele Kurven werden benötigt?

$$n \approx \frac{28}{\rho^2} \propto \frac{1}{\mathrm{SNR}} = \frac{\mathrm{Var}(P_{Noise})}{\mathrm{Var}(P_{Signal})}$$

Wie groß ist ρ bei falscher Vorhersage?

Hinweise zu Angriffen mit Korrelations-Koeffizient

Wie hoch ist der Korrelations-Koeffizient?

$$\rho = \sqrt{\frac{\text{SNR}}{\text{SNR}+1}} = \sqrt{\frac{1}{1+\text{SNR}^{-1}}}$$

Wie viele Kurven werden benötigt?

$$n \approx \frac{28}{\rho^2} \propto \frac{1}{\text{SNR}} = \frac{\text{Var}(P_{Noise})}{\text{Var}(P_{Signal})}$$

Wie groß ist ρ bei falscher Vorhersage?

Im Intervall $\pm\frac{4}{\sqrt{n}}$

CPA - Correlation Power Analysis

CPA - Correlation Power Analysis

- Typischer Angriff für Power Analysen

CPA - Correlation Power Analysis

- Typischer Angriff für Power Analysen
- Grundsätzlich wie DPA, nutzt jedoch mehr Informationen

CPA - Correlation Power Analysis

Vorgehen:

1. Messkurven & I/O Aufzeichnen (wie DPA)

CPA - Correlation Power Analysis

Vorgehen:

1. Messkurven & I/O Aufzeichnen (wie DPA)
2. Ggf. Preprocessing (Alignment und Kompression)

CPA - Correlation Power Analysis

Vorgehen:

1. Messkurven & I/O Aufzeichnen (wie DPA)
2. Ggf. Preprocessing (Alignment und Kompression)
3. Zwischenwerte unter Annahme eines SubKeys vorhersagen

CPA - Correlation Power Analysis

Vorgehen:

1. Messkurven & I/O Aufzeichnen (wie DPA)
2. Ggf. Preprocessing (Alignment und Kompression)
3. Zwischenwerte unter Annahme eines SubKeys vorhersagen
4. Power Model verwenden um die vorhergesagten Zwischenwerte auf einen Stromverbrauch abzubilden (typisch: HW Model)

CPA - Correlation Power Analysis

Vorgehen:

1. Messkurven & I/O Aufzeichnen (wie DPA)
2. Ggf. Preprocessing (Alignment und Kompression)
3. Zwischenwerte unter Annahme eines SubKeys vorhersagen
4. Power Model verwenden um die vorhergesagten Zwischenwerte auf einen Stromverbrauch abzubilden (typisch: HW Model)
5. Für jeden SubKey Hypothesen mit Messwerten vergleichen (Korrelations-Koeffizient)

CPA - Correlation Power Analysis

Vorgehen:

1. Messkurven & I/O Aufzeichnen (wie DPA)
2. Ggf. Preprocessing (Alignment und Kompression)
3. Zwischenwerte unter Annahme eines SubKeys vorhersagen
4. Power Model verwenden um die vorhergesagten Zwischenwerte auf einen Stromverbrauch abzubilden (typisch: HW Model)
5. Für jeden SubKey Hypothesen mit Messwerten vergleichen (Korrelations-Koeffizient)
6. Schlüssel Identifizieren

CPA - Correlation Power Analysis

Beispiel: Stromkurve AES (MATLAB)

Summary

- Seitenkanäle

Summary

- Seitenkanäle
- SPA

Summary

- Seitenkanäle
- SPA
- DPA

Summary

- Seitenkanäle
- SPA
- DPA
- Stromverbrauch und Statistik

Summary

- Seitenkanäle
- SPA
- DPA
- Stromverbrauch und Statistik
- CPA

Outlook

- Gegenmaßnahmen

Outlook

- Gegenmaßnahmen
- Angriffe auf Gegenmaßnahmen

Outlook

- Gegenmaßnahmen
- Angriffe auf Gegenmaßnahmen
- Template Angriffe

Outlook

- Gegenmaßnahmen
- Angriffe auf Gegenmaßnahmen
- Template Angriffe
- Fault Injection

Outlook

- Gegenmaßnahmen
- Angriffe auf Gegenmaßnahmen
- Template Angriffe
- Fault Injection
- Mutual Information Analysis

Outlook

- Gegenmaßnahmen
- Angriffe auf Gegenmaßnahmen
- Template Angriffe
- Fault Injection
- Mutual Information Analysis
- Principal Component Analysis und Local Discriminant Analysis

Thanks For Your Attention!

Any Questions?

Vorsicht Hintertür!

Seitenkanäle als versteckte Systemschwachstellen

Markus Kasper

Lehrstuhl für Embedded Security
Ruhr-Universität Bochum

Labor Vortrag 2010

ToDo:

Feedback ins Wiki posten!

Dem Labor etwas spenden damit wir ein brauchbares Speicheroszilloskop kaufen können!