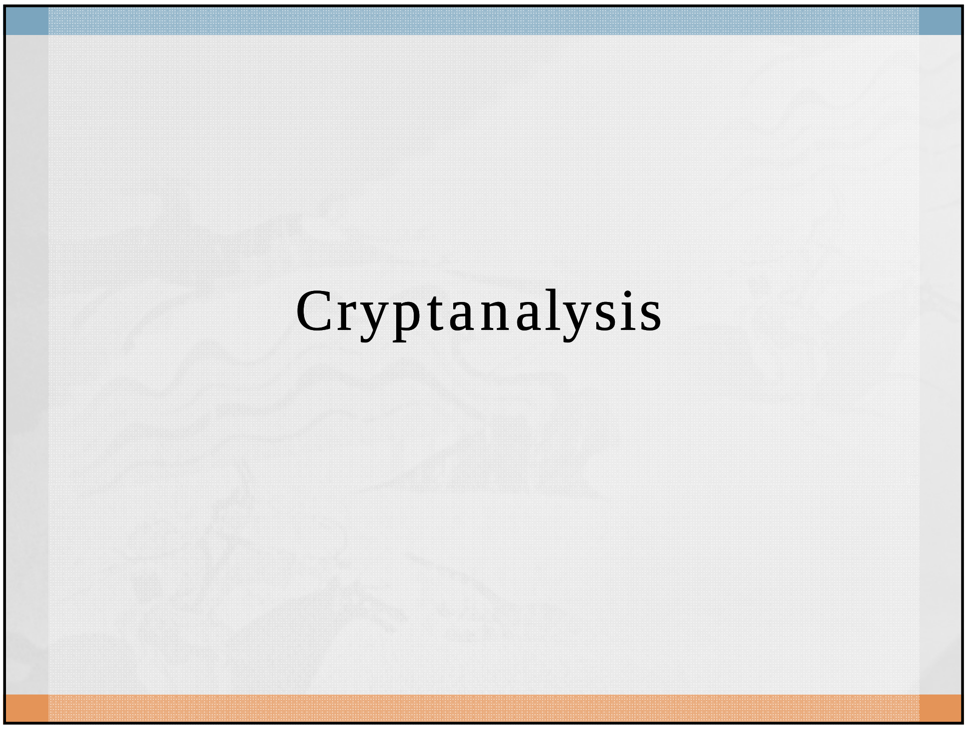
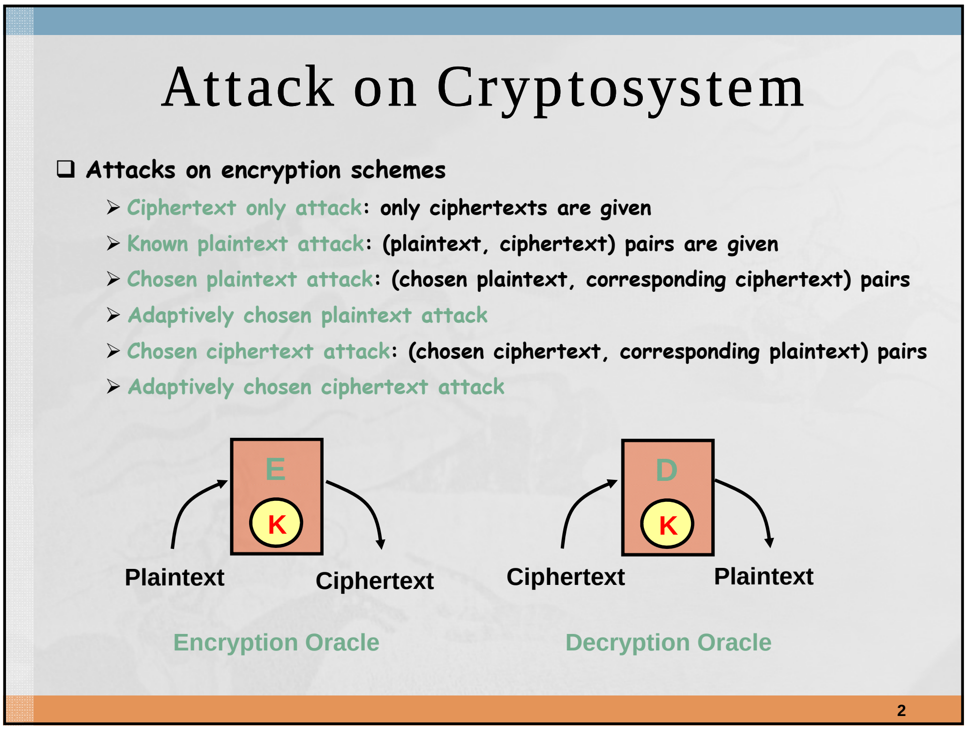


Cryptanalysis

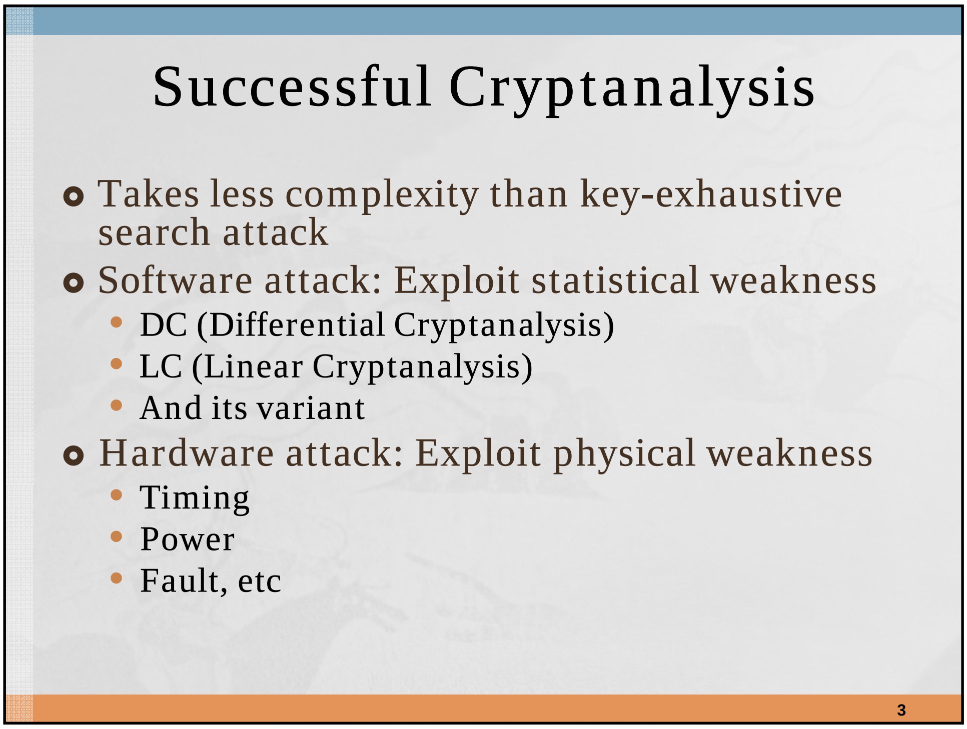


Attack on Cryptosystem

- ❑ Attacks on encryption schemes
  - ➢ Ciphertext only attack: only ciphertexts are given
  - ➢ Known plaintext attack: (plaintext, ciphertext) pairs are given
  - ➢ Chosen plaintext attack: (chosen plaintext, corresponding ciphertext) pairs
  - ➢ Adaptively chosen plaintext attack
  - ➢ Chosen ciphertext attack: (chosen ciphertext, corresponding plaintext) pairs
  - ➢ Adaptively chosen ciphertext attack

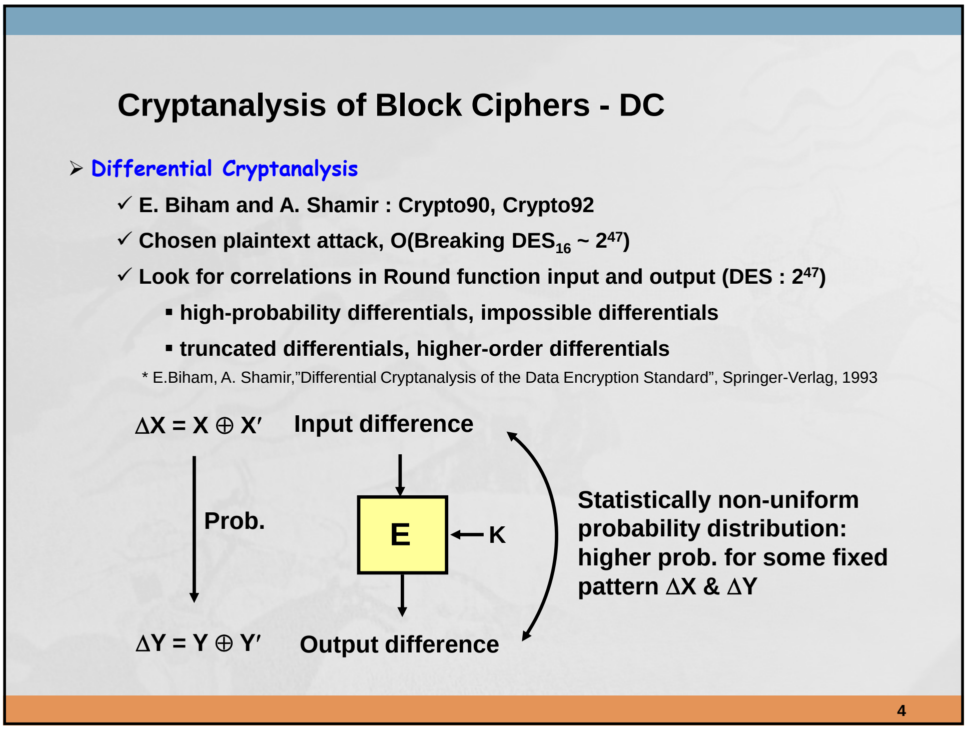
Encryption Oracle

Decryption Oracle



Successful Cryptanalysis

- Takes less complexity than key-exhaustive search attack
- Software attack: Exploit statistical weakness
  - DC (Differential Cryptanalysis)
  - LC (Linear Cryptanalysis)
  - And its variant
- Hardware attack: Exploit physical weakness
  - Timing
  - Power
  - Fault, etc



Cryptanalysis of Block Ciphers - DC

➢ Differential Cryptanalysis

- ✓ E. Biham and A. Shamir : Crypto90, Crypto92
- ✓ Chosen plaintext attack, O(Breaking DES$_{16}$ ~ 2^{47})
- ✓ Look for correlations in Round function input and output (DES : 2^{47})
  - ▪ high-probability differentials, impossible differentials
  - ▪ truncated differentials, higher-order differentials

* E.Biham, A. Shamir,"Differential Cryptanalysis of the Data Encryption Standard", Springer-Verlag, 1993

$\Delta X = X \oplus X'$ Input difference

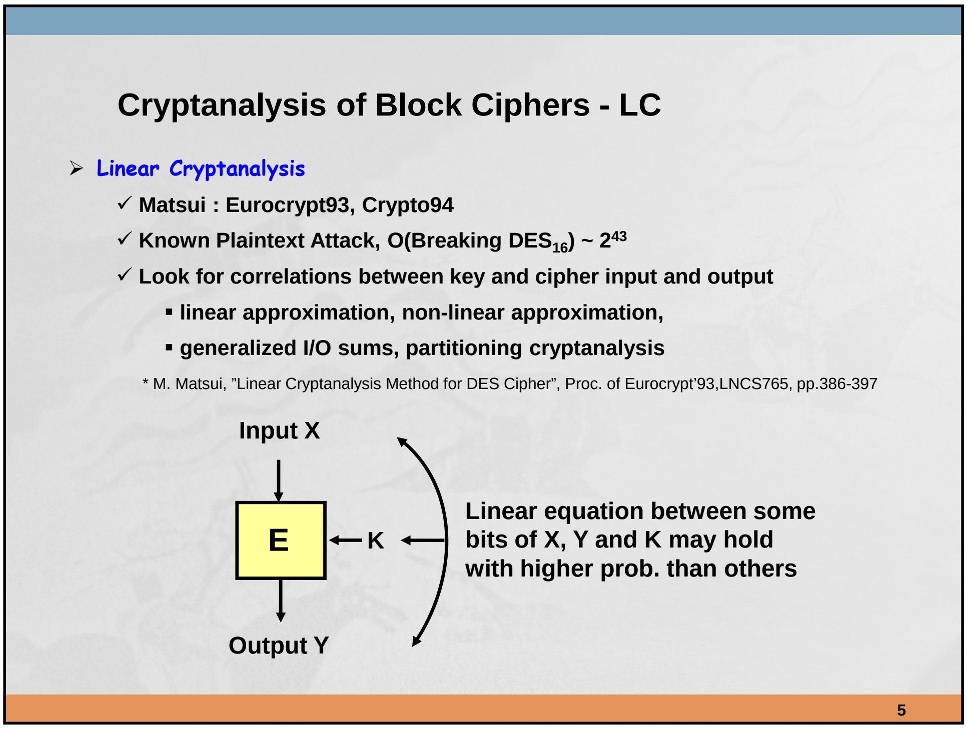
Prob.

E

K

$\Delta Y = Y \oplus Y'$ Output difference

Statistically non-uniform probability distribution: higher prob. for some fixed pattern ΔX & ΔY



Cryptanalysis of Block Ciphers - LC

- Linear Cryptanalysis
  - ✓ Matsui : Eurocrypt93, Crypto94
  - ✓ Known Plaintext Attack, O(Breaking DES_{16}) ~ 2^{43}
  - ✓ Look for correlations between key and cipher input and output
    - linear approximation, non-linear approximation,
    - generalized I/O sums, partitioning cryptanalysis

* M. Matsui, "Linear Cryptanalysis Method for DES Cipher", Proc. of Eurocrypt'93,LNCS765, pp.386-397

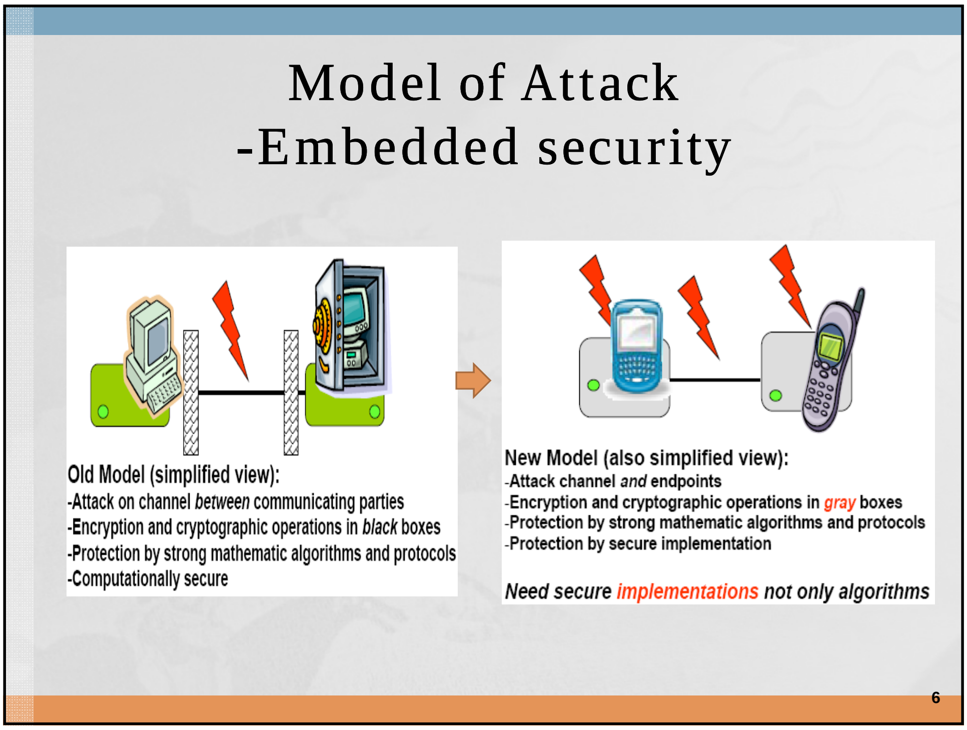
Input X

E

K

Output Y

Linear equation between some bits of X, Y and K may hold with higher prob. than others



Model of Attack -Embedded security

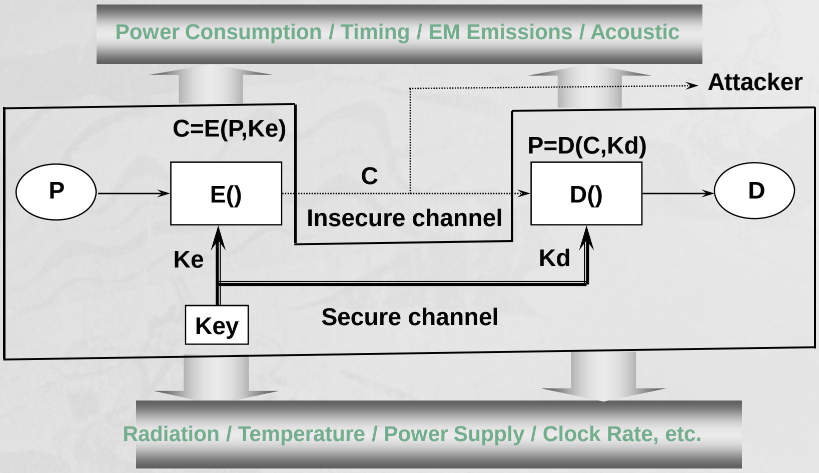
Old Model (simplified view):
-Attack on channel *between* communicating parties
-Encryption and cryptographic operations in *black* boxes
-Protection by strong mathematic algorithms and protocols
-Computationally secure

New Model (also simplified view):
-Attack channel *and* endpoints
-Encryption and cryptographic operations in *gray* boxes
-Protection by strong mathematic algorithms and protocols
-Protection by secure implementation

Need secure implementations not only algorithms

Side Channel

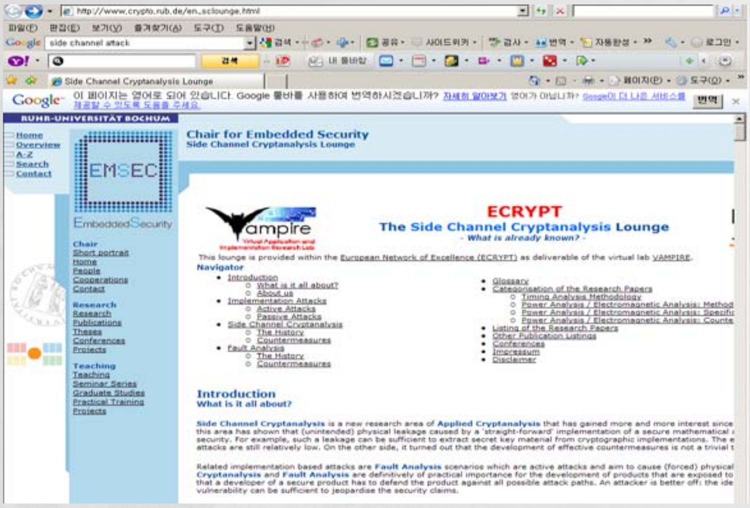
- Traditional Cryptographic Model vs. Side Channel



Power Consumption / Timing / EM Emissions / Acoustic
Attacker
C=E(P,Ke)
P=D(C,Kd)
P
E()
C
Insecure channel
D()
D
Ke
Kd
Key
Secure channel
Radiation / Temperature / Power Supply / Clock Rate, etc.

Let's watch video on SCA http://www.cryptography.com/technology/dpa.html

SCA Crypto Lounge



Chair for Embedded Security
Side Channel Cryptanalysis Lounge

ECRYPT
The Side Channel Cryptanalysis Lounge
- What is already known? -

This lounge is provided within the European Network of Excellence (ECRYPT) as deliverable of the virtual lab VAMPIRE.

Navigator

- Introduction
  - What is it all about?
  - About us
- Implementation Attacks
  - Active Attacks
  - Passive Attacks
- Side Channel Cryptanalysis
  - The History
  - Countermeasures
- Fault Analysis
  - The History
  - Countermeasures

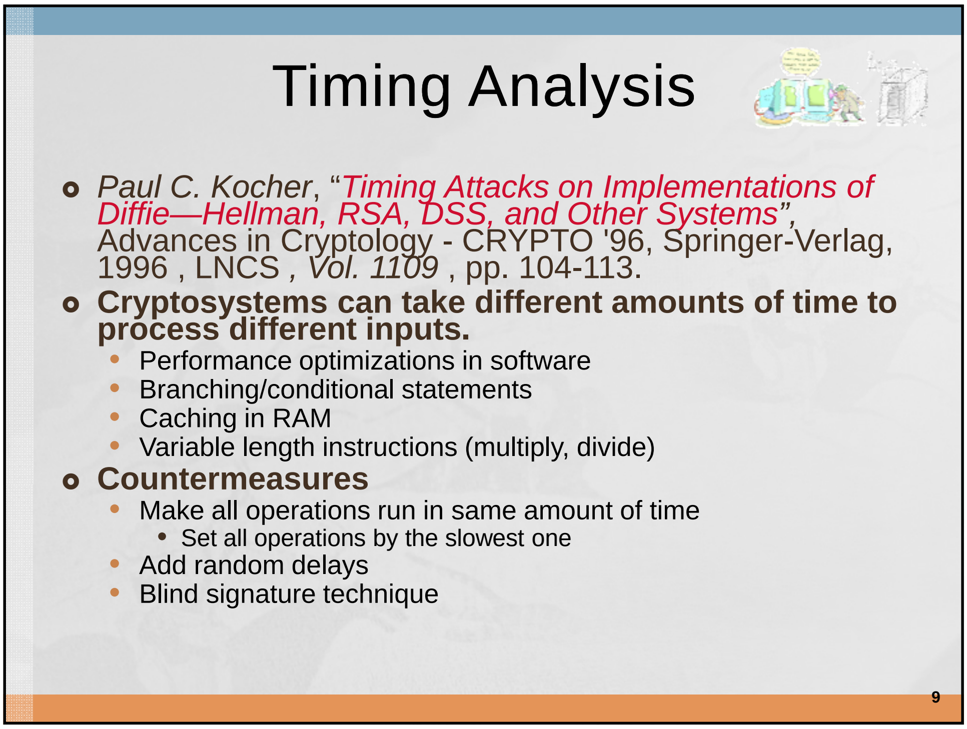
- Glossary
- Categorisation of the Research Papers
  - Timing Analysis Methodology
  - Power Analysis / Electromagnetic Analysis: Method
  - Power Analysis / Electromagnetic Analysis: Specific
  - Power Analysis / Electromagnetic Analysis: Counte
- Listing of the Research Papers
- Other Publication Listings
- Conferences
- Impressum
- Disclaimer

Introduction
What is it all about?

Side Channel Cryptanalysis is a new research area of **Applied Cryptanalysis** that has gained more and more interest since this area has shown that (unintended) physical leakage caused by a 'straight-forward' implementation of a secure mathematical security. For example, such a leakage can be sufficient to extract secret key material from cryptographic implementations. The e attacks are still relatively low. On the other side, it turned out that the development of effective countermeasures is not a trivial t

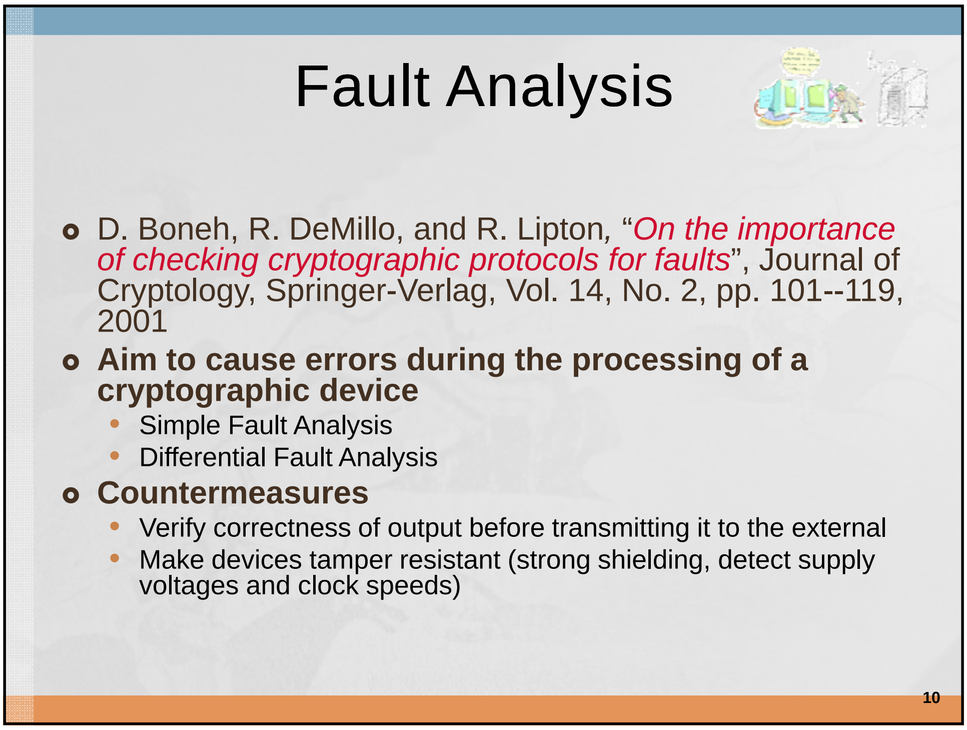
Related implementation based attacks are **Fault Analysis** scenarios which are active attacks and aim to cause (forced) physical **Cryptanalysis** and **Fault Analysis** are definitively of practical importance for the development of products that are exposed to that a developer of a secure product has to defend the product against all possible attack paths. An attacker is better off: the ide vulnerability can be sufficient to jeopardise the security claims.

Click for details



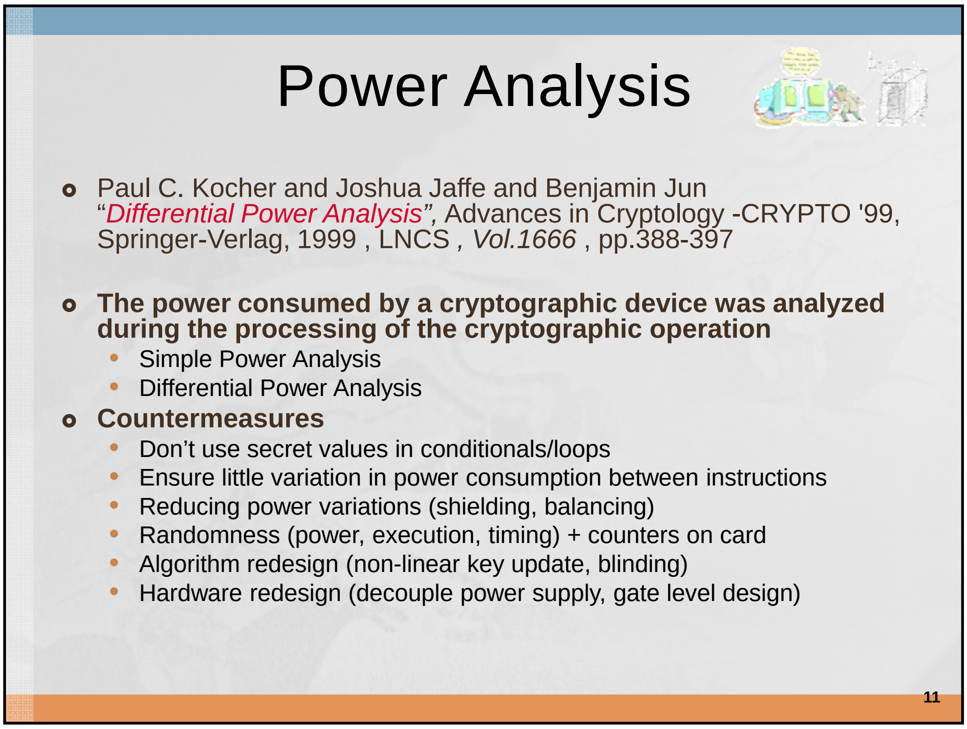
Timing Analysis

- *Paul C. Kocher*, "*Timing Attacks on Implementations of Diffie—Hellman, RSA, DSS, and Other Systems*", Advances in Cryptology - CRYPTO '96, Springer-Verlag, 1996 , LNCS , *Vol. 1109* , pp. 104-113.
- **Cryptosystems can take different amounts of time to process different inputs.**
  - Performance optimizations in software
  - Branching/conditional statements
  - Caching in RAM
  - Variable length instructions (multiply, divide)
- **Countermeasures**
  - Make all operations run in same amount of time
    - Set all operations by the slowest one
  - Add random delays
  - Blind signature technique



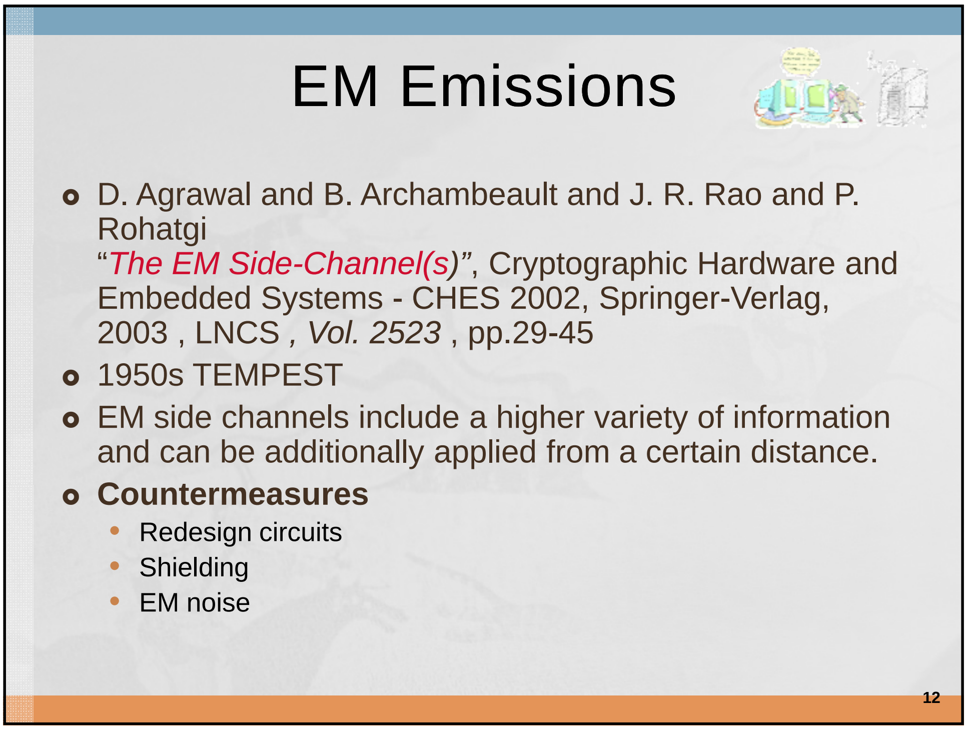
Fault Analysis

- D. Boneh, R. DeMillo, and R. Lipton, "*On the importance of checking cryptographic protocols for faults*", Journal of Cryptology, Springer-Verlag, Vol. 14, No. 2, pp. 101--119, 2001
- **Aim to cause errors during the processing of a cryptographic device**
  - Simple Fault Analysis
  - Differential Fault Analysis
- **Countermeasures**
  - Verify correctness of output before transmitting it to the external
  - Make devices tamper resistant (strong shielding, detect supply voltages and clock speeds)



Power Analysis

- Paul C. Kocher and Joshua Jaffe and Benjamin Jun
  "*Differential Power Analysis*", Advances in Cryptology -CRYPTO '99, Springer-Verlag, 1999 , LNCS , *Vol.1666* , pp.388-397
- **The power consumed by a cryptographic device was analyzed during the processing of the cryptographic operation**
  - Simple Power Analysis
  - Differential Power Analysis
- **Countermeasures**
  - Don't use secret values in conditionals/loops
  - Ensure little variation in power consumption between instructions
  - Reducing power variations (shielding, balancing)
  - Randomness (power, execution, timing) + counters on card
  - Algorithm redesign (non-linear key update, blinding)
  - Hardware redesign (decouple power supply, gate level design)



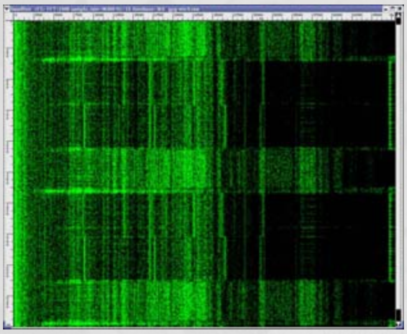
EM Emissions

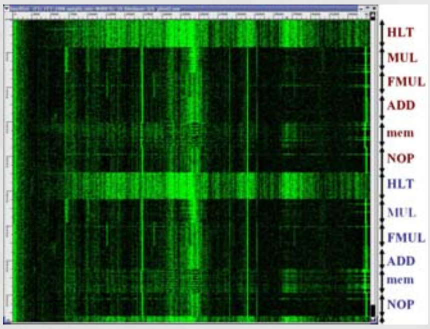
- D. Agrawal and B. Archambeault and J. R. Rao and P. Rohatgi
  "*The EM Side-Channel(s)*", Cryptographic Hardware and Embedded Systems - CHES 2002, Springer-Verlag, 2003 , LNCS , *Vol. 2523* , pp.29-45
- 1950s TEMPEST
- EM side channels include a higher variety of information and can be additionally applied from a certain distance.
- **Countermeasures**
  - Redesign circuits
  - Shielding
  - EM noise

Acoustic Analysis

Acoustic Analysis

- *Keyboard Acoustic Emanations*, Dmitri Asonov and Rakesh Agrawal, IBM Almaden Research Center, 2004.
- Acoustic cryptanalysis - On noisy people and noisy machines by Adi Shamir and Eran Tromer



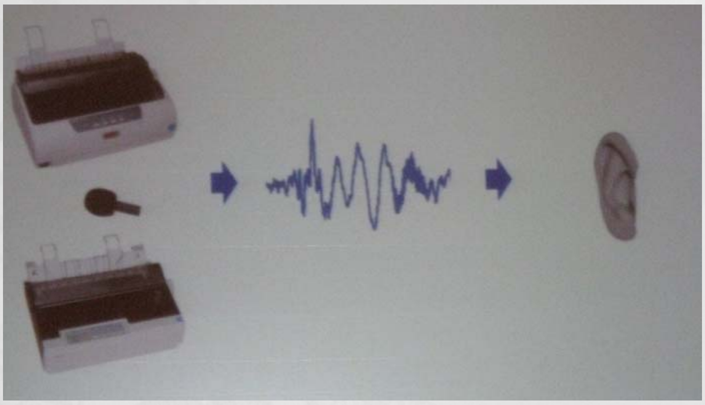


ACOUSTIC SIDE-CHANNEL ATTACKS ON PRINTERS

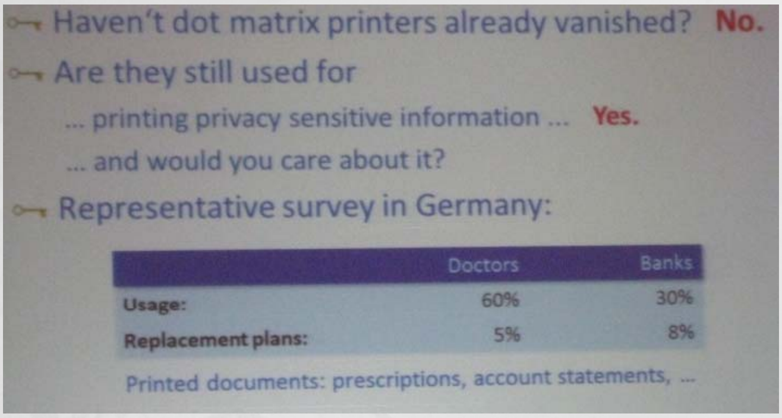
Michael Backes, *Saarland University and Max Planck Institute for Software Systems (MPI-SWS);* Markus Dürmuth, Sebastian Gerling, Manfred Pinkal, and Caroline Sporleder,

Saarland University , 19th USENIX Security Symposium, Washington DC, Aug. 11-13, 2010

Motivation



Why should you care?

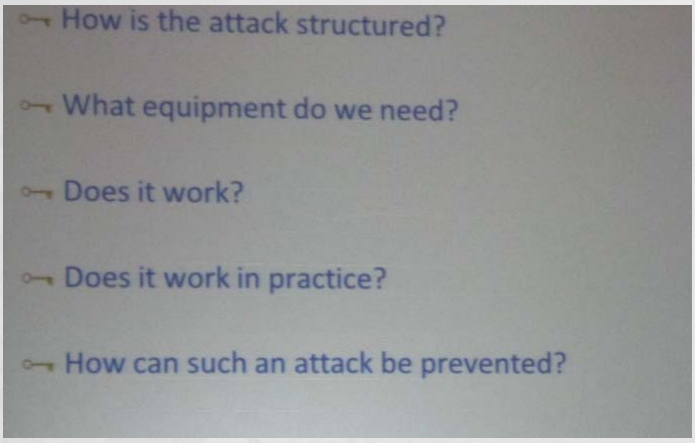


- Haven't dot matrix printers already vanished? **No.**
- Are they still used for
  ... printing privacy sensitive information ... **Yes.**
  ... and would you care about it?
- Representative survey in Germany:

	Doctors	Banks
Usage:	60%	30%
Replacement plans:	5%	8%

Printed documents: prescriptions, account statements, ...

How does our attack work

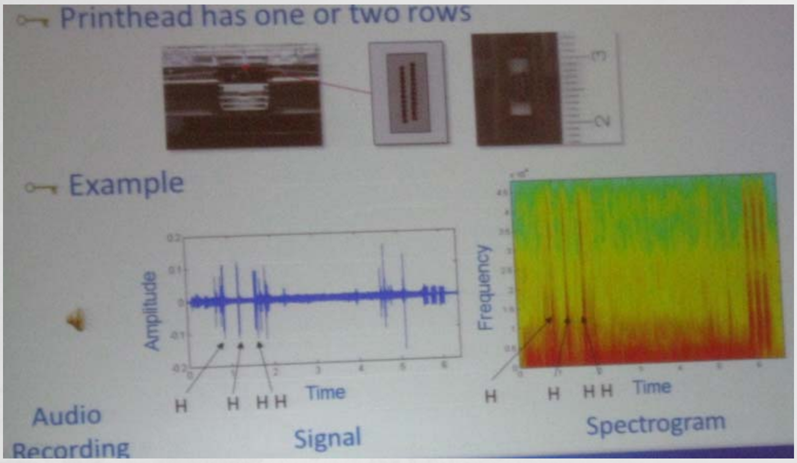


- How is the attack structured?
- What equipment do we need?
- Does it work?
- Does it work in practice?
- How can such an attack be prevented?

Acoustic Cryptanalysis

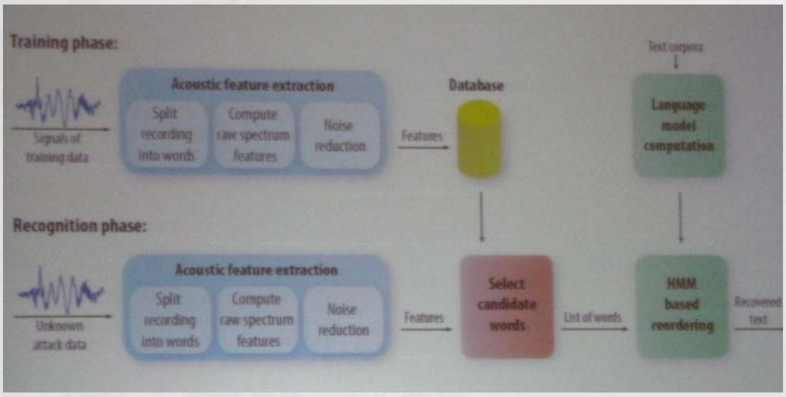
D. Asonov and R. Agrawal. Keyboard Acoustic Emanations, 2004.
L. Zhuang, F. Zhou, and J.D. Tygar. Keyboard Acoustic Emanations Revisited, 2005.
Y. Berger, A. Wool, and A. Yeredor. Dictionary attacks using keyboard acoustic emanations, 2006.
A. Shamir and E. Tromer. Acoustic cryptanalysis: On nosy people and noisy machines, 2004.
R. Briol. Emanation: How to keep your data confidential, 1991.

Dot Matrix Printer-Details



- Printhead has one or two rows
- Example

How it works



Training phase:

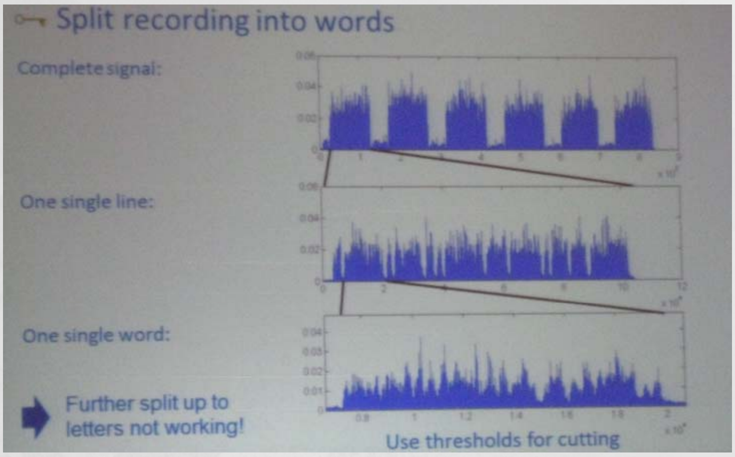
Signals of training data → Acoustic feature extraction: Split recording into words | Compute raw spectrum features | Noise reduction → Features → Database

Text corpora → Language model computation

Recognition phase:

Unknown attack data → Acoustic feature extraction: Split recording into words | Compute raw spectrum features | Noise reduction → Features → Select candidate words → List of words → HMM based reordering → Recovered text

Acoustic Feature Extraction(1/3)



Split recording into words

Complete signal:

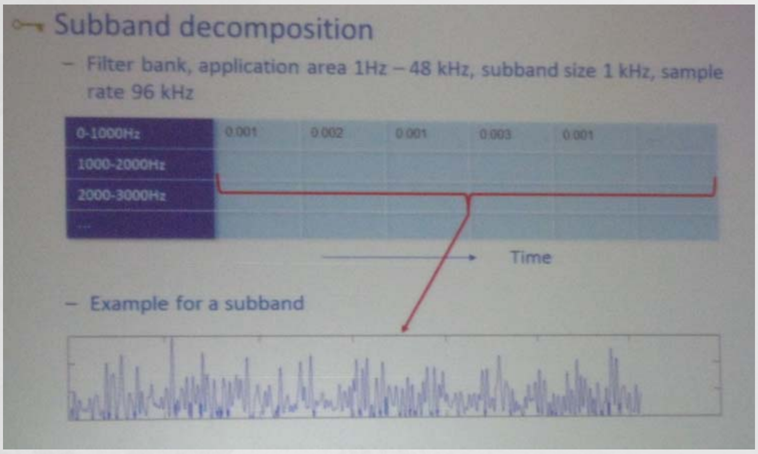
One single line:

One single word:

Further split up to letters not working!

Use thresholds for cutting

Acoustic Feature Extraction(2/3)



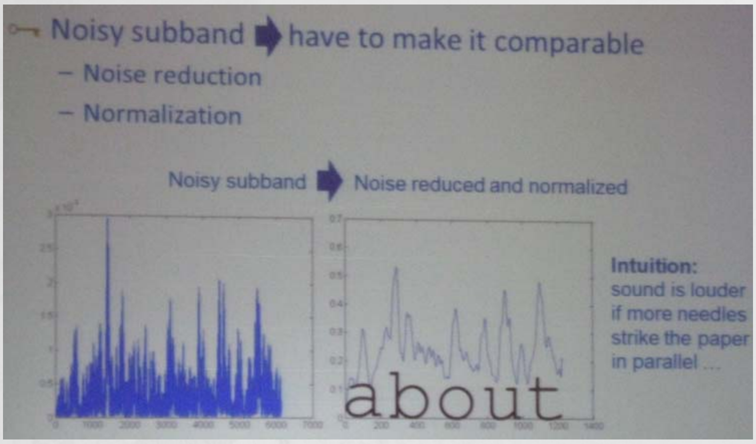
Subband decomposition

- Filter bank, application area 1Hz – 48 kHz, subband size 1 kHz, sample rate 96 kHz

0-1000Hz	0.001	0.002	0.001	0.003	0.001	
1000-2000Hz						
2000-3000Hz						
...						

- Example for a subband

Acoustic Feature Extraction(3/3)

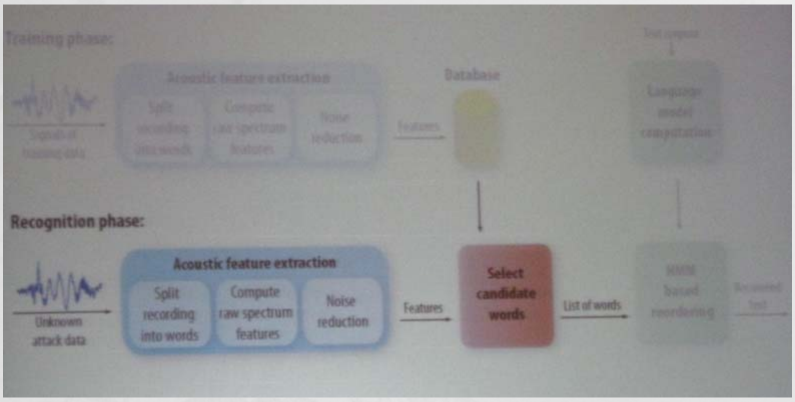


- Noisy subband ➡ have to make it comparable
  - Noise reduction
  - Normalization

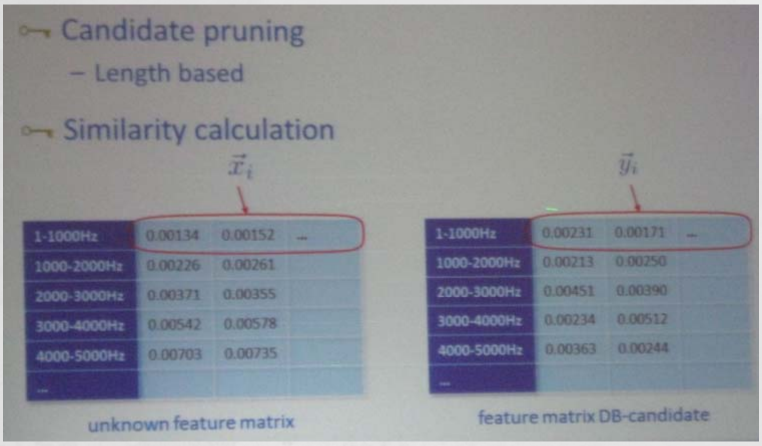
Noisy subband ➡ Noise reduced and normalized

Intuition:
sound is louder if more needles strike the paper in parallel …

Recognition Phase



Select Candidate Words(1/2)



- Candidate pruning
  - Length based
- Similarity calculation

$\vec{x}_i$

1-1000Hz	0.00134	0.00152	...
1000-2000Hz	0.00226	0.00261	
2000-3000Hz	0.00371	0.00355	
3000-4000Hz	0.00542	0.00578	
4000-5000Hz	0.00703	0.00735	
...			

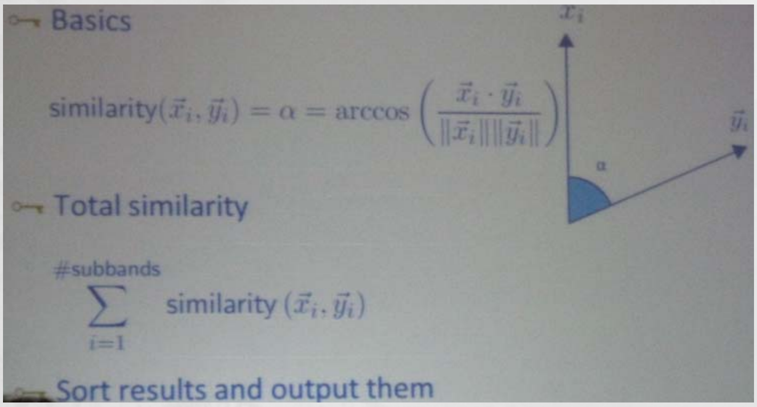
unknown feature matrix

$\vec{y}_i$

1-1000Hz	0.00231	0.00171	...
1000-2000Hz	0.00213	0.00250	
2000-3000Hz	0.00451	0.00390	
3000-4000Hz	0.00234	0.00512	
4000-5000Hz	0.00363	0.00244	
...			

feature matrix DB-candidate

Select Candidate Words(1/2)



- Basics

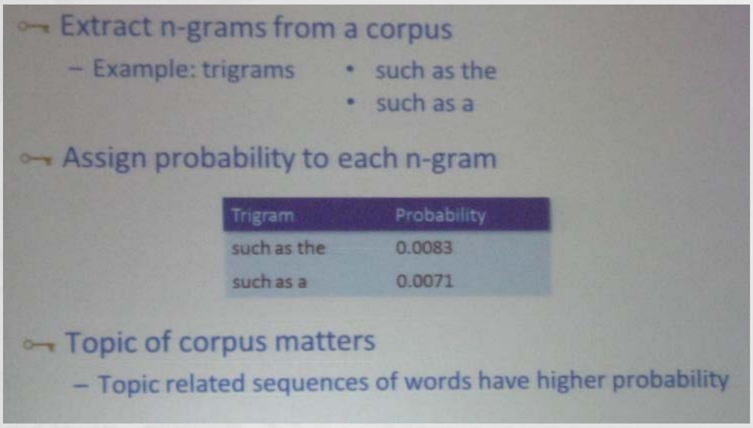
$$\text{similarity}(\vec{x}_i, \vec{y}_i) = \alpha = \arccos\left(\frac{\vec{x}_i \cdot \vec{y}_i}{\|\vec{x}_i\|\|\vec{y}_i\|}\right)$$

- Total similarity

$$\sum_{i=1}^{\#\text{subbands}} \text{similarity}\,(\vec{x}_i, \vec{y}_i)$$

- Sort results and output them

Language Model Computation

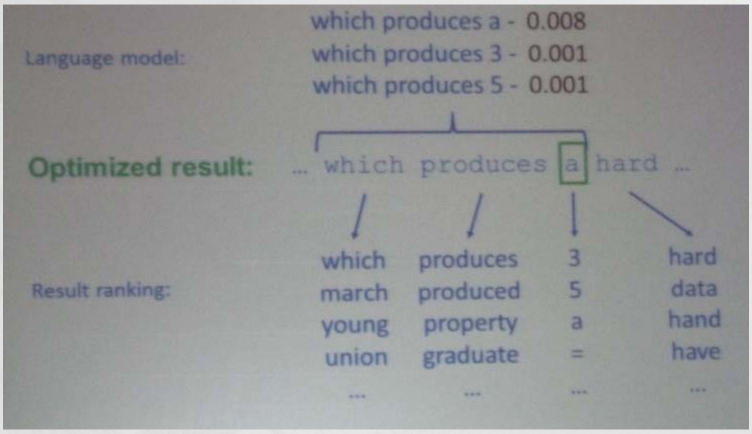


- Extract n-grams from a corpus
  - Example: trigrams
    - such as the
    - such as a
- Assign probability to each n-gram

Trigram	Probability
such as the	0.0083
such as a	0.0071

- Topic of corpus matters
  - Topic related sequences of words have higher probability

Reordering based on Hidden Markov Model



Language model:

which produces a - 0.008
which produces 3 - 0.001
which produces 5 - 0.001

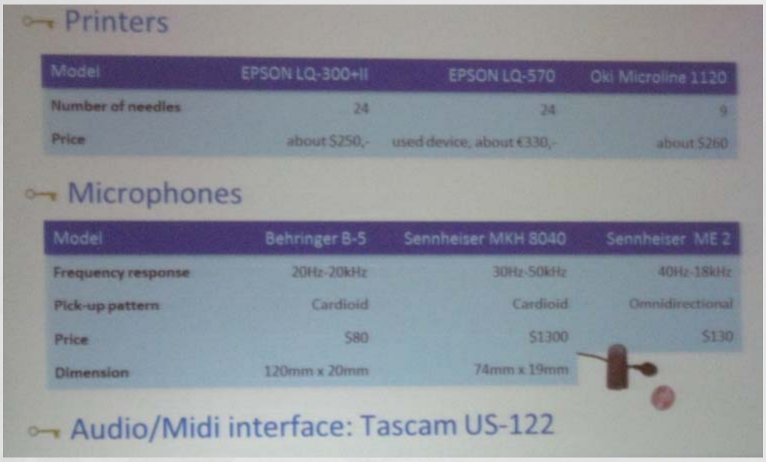
Optimized result: ... which produces a hard ...

Result ranking:

which	produces	3	hard
march	produced	5	data
young	property	a	hand
union	graduate	=	have
...	...	...	...

What equipment do we need?

Equipment - Hardware



- Printers

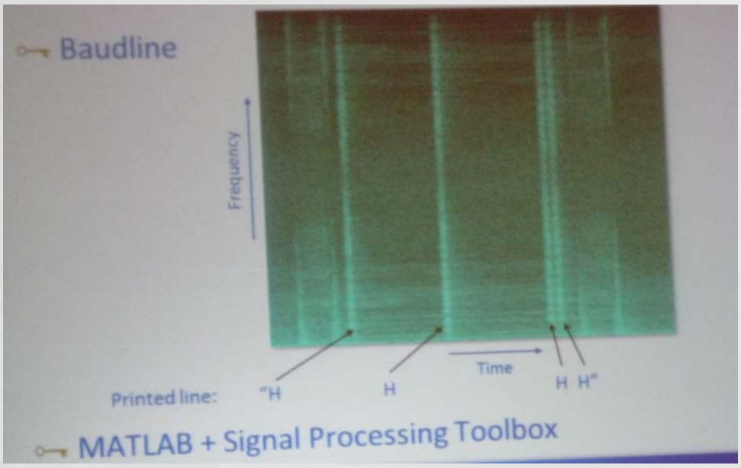
Model	EPSON LQ-300+II	EPSON LQ-570	Oki Microline 1120
Number of needles	24	24	9
Price	about $250,-	used device, about €330,-	about $260

- Microphones

Model	Behringer B-5	Sennheiser MKH 8040	Sennheiser ME 2
Frequency response	20Hz-20kHz	30Hz-50kHz	40Hz-18kHz
Pick-up pattern	Cardioid	Cardioid	Omnidirectional
Price	$80	$1300	$130
Dimension	120mm x 20mm	74mm x 19mm	

- Audio/Midi interface: Tascam US-122

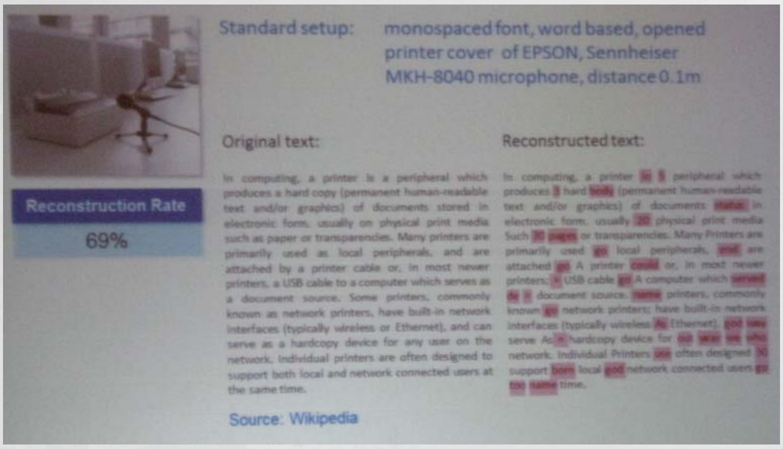
Equipment - Software



- Baudline

- MATLAB + Signal Processing Toolbox

Results(1/2)



Standard setup: monospaced font, word based, opened printer cover of EPSON, Sennheiser MKH-8040 microphone, distance 0.1m

Reconstruction Rate

69%

Original text:

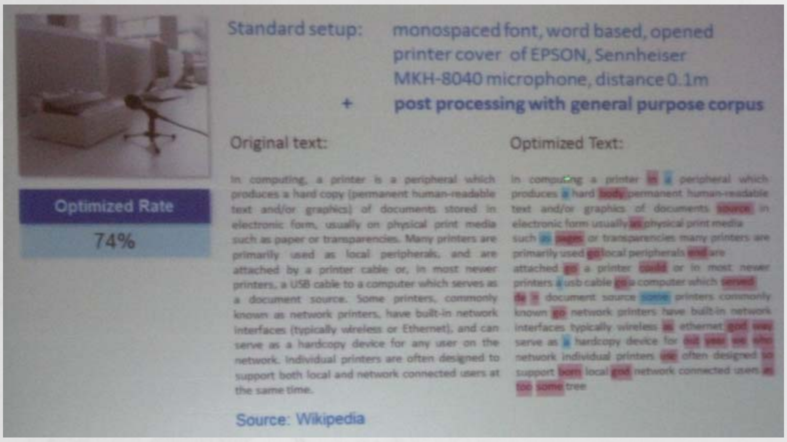
In computing, a printer is a peripheral which produces a hard copy (permanent human-readable text and/or graphics) of documents stored in electronic form, usually on physical print media such as paper or transparencies. Many printers are primarily used as local peripherals, and are attached by a printer cable or, in most newer printers, a USB cable to a computer which serves as a document source. Some printers, commonly known as network printers, have built-in network interfaces (typically wireless or Ethernet), and can serve as a hardcopy device for any user on the network. Individual printers are often designed to support both local and network connected users at the same time.

Source: Wikipedia

Reconstructed text:

In computing, a printer is 5 peripheral which produces 3 hard body (permanent human-readable text and/or graphics) of documents status in electronic form. usually 20 physical print media Such 30 pages or transparencies. Many Printers are primarily used go local peripherals, and are attached go A printer could or, in most newer printers; a USB cable go A computer which served As a document source. Same printers, commonly known go network printers; have built-in network interfaces (typically wireless As Ethernet). god say serve As a hardcopy device for our year we who network. Individual Printers use often designed 30 support born local god network connected users go too name time.

Results(2/2)



Standard setup: monospaced font, word based, opened printer cover of EPSON, Sennheiser MKH-8040 microphone, distance 0.1m

+ **post processing with general purpose corpus**

Optimized Rate

74%

Original text:

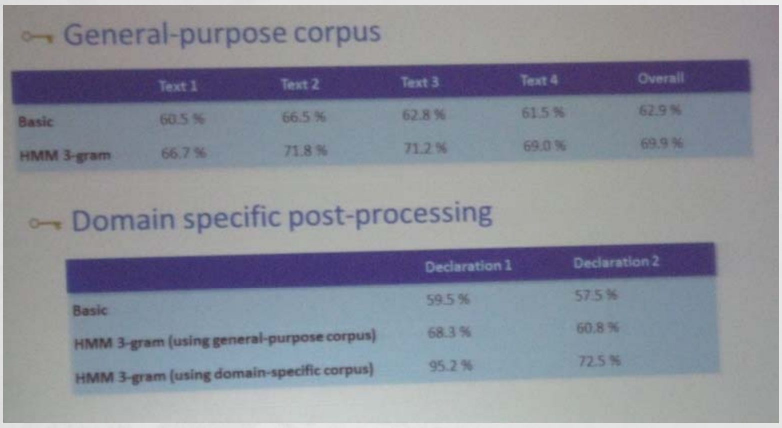
In computing, a printer is a peripheral which produces a hard copy (permanent human-readable text and/or graphics) of documents stored in electronic form, usually on physical print media such as paper or transparencies. Many printers are primarily used as local peripherals, and are attached by a printer cable or, in most newer printers, a USB cable to a computer which serves as a document source. Some printers, commonly known as network printers, have built-in network interfaces (typically wireless or Ethernet), and can serve as a hardcopy device for any user on the network. Individual printers are often designed to support both local and network connected users at the same time.

Optimized Text:

In computing a printer in a peripheral which produces a hard body permanent human-readable text and/or graphics of documents source in electronic form usually as physical print media such as pages or transparencies many printers are primarily used go local peripherals end are attached go a printer could or in most newer printers a usb cable go a computer which served de a document source some printers commonly known go network printers have built-in network interfaces typically wireless as ethernet god way serve as a hardcopy device for but year we who network individual printers are often designed so support born local god network connected users at too some tree

Source: Wikipedia

Results (Standard Setup)



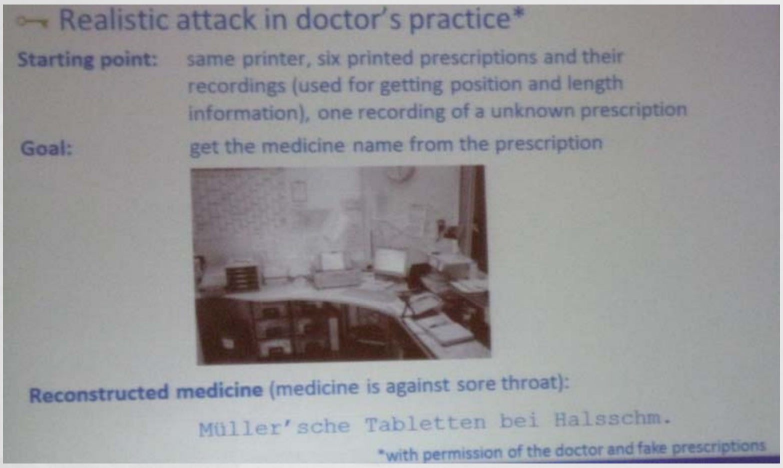
General-purpose corpus

	Text 1	Text 2	Text 3	Text 4	Overall
Basic	60.5 %	66.5 %	62.8 %	61.5 %	62.9 %
HMM 3-gram	66.7 %	71.8 %	71.2 %	69.0 %	69.9 %

Domain specific post-processing

	Declaration 1	Declaration 2
Basic	59.5 %	57.5 %
HMM 3-gram (using general-purpose corpus)	68.3 %	60.8 %
HMM 3-gram (using domain-specific corpus)	95.2 %	72.5 %

In-field Attack



- Realistic attack in doctor's practice*

Starting point: same printer, six printed prescriptions and their recordings (used for getting position and length information), one recording of a unknown prescription

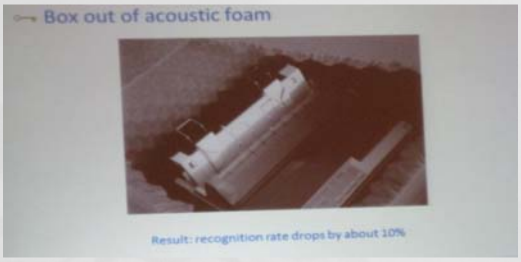
Goal: get the medicine name from the prescription

Reconstructed medicine (medicine is against sore throat):

`Müller'sche Tabletten bei Halsschm.`

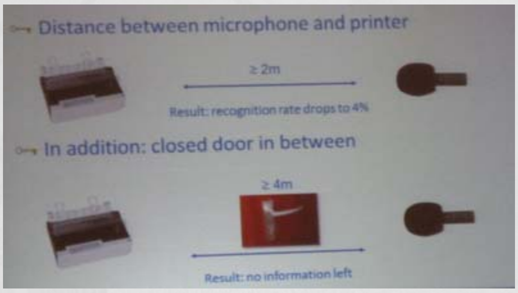
*with permission of the doctor and fake prescriptions

Countermeasures



- Box out of acoustic foam

Result: recognition rate drops by about 10%



- Distance between microphone and printer

≥ 2m

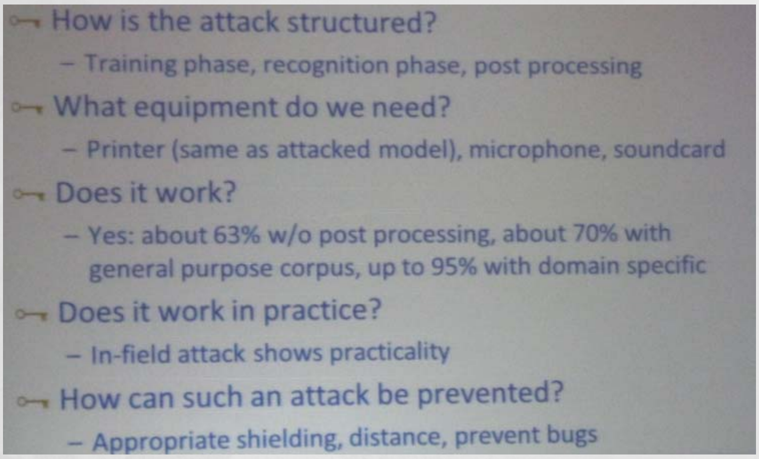
Result: recognition rate drops to 4%

- In addition: closed door in between

≥ 4m

Result: no information left

Conclusion



- How is the attack structured?
  - Training phase, recognition phase, post processing
- What equipment do we need?
  - Printer (same as attacked model), microphone, soundcard
- Does it work?
  - Yes: about 63% w/o post processing, about 70% with general purpose corpus, up to 95% with domain specific
- Does it work in practice?
  - In-field attack shows practicality
- How can such an attack be prevented?
  - Appropriate shielding, distance, prevent bugs