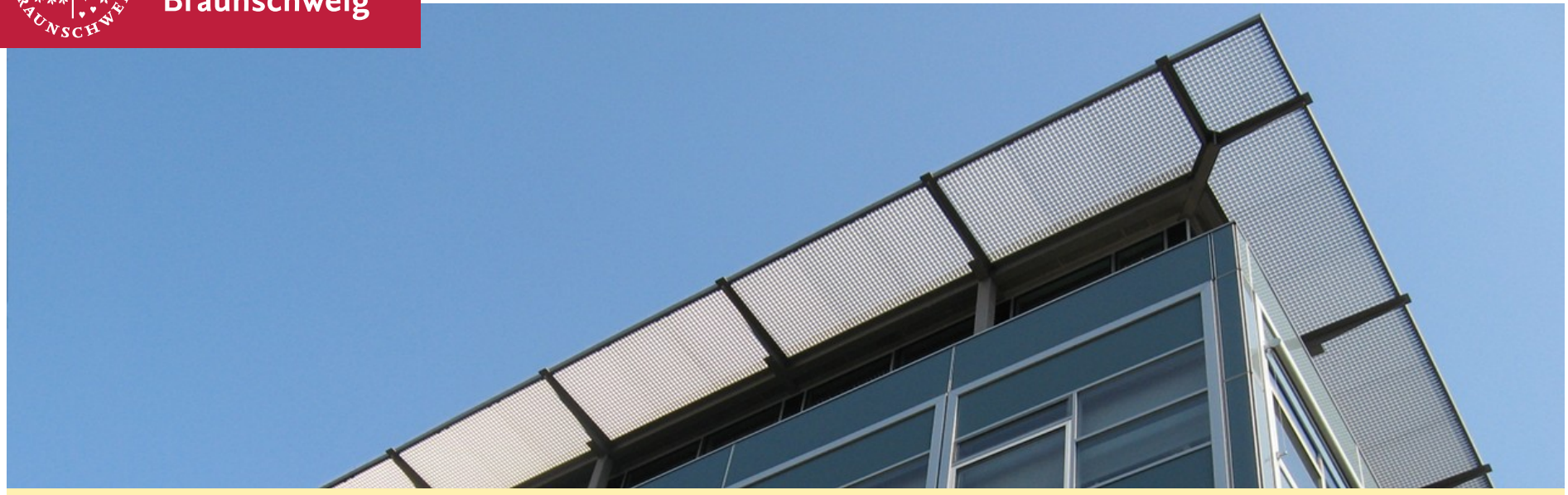
Technische Universität Braunschweig

Institute of Operating Systems and Computer Networks



Secure communication based on noisy input data

Fuzzy Commitment schemes

Stephan Sigg

May 24, 2011

Overview and Structure

05.04.2011 Organisational
15.04.2011 Introduction
19.04.2011 Classification methods (Basic recognition, Bayesian, Non-parametric)
26.04.2011 Classification methods (Linear discriminant, Neural networks)
03.05.2011 Classification methods (Sequential, Stochastic)
10.05.2011 Feature extraction from audio data
17.05.2011 Feature extraction from the RF channel
24.05.2011 Fuzzy Commitment
31.05.2011 Fuzzy Extractors
07.06.2011 Error correcting codes
21.06.2011 Entropy
28.06.2011 Physically unclonable functions

Outline

Fuzzy cryptography

History of biometrics

Egyptian times: Early use of biometrics (physiological properties to distinguish traders)

14th century: Chinese merchants used hand palm prints and footprints on paper to distinguish young children

19th century: Biometric methods used to solve crime (head and body measurements to identify convicted criminals)

End of 19th century: Fingerprints became popular for forensic use

The system had the problem, that no easy way of sorting and identifying fingerprints was known

$\approx$ 1900: Classification system distinguishing fingerprint classes proposed

Variations of this system are the basis of many fingerprint identification systems nowadays

Fuzzy cryptography

History of biometrics

The face of indivuduals has long be the modality of choice for official documents

- Matching was commonly accomplished by the human eye
- Automated systems for face recognition are increasingly used
- Faces characterised by landmark points (corner of eyebrows, tip of nose) or local texture of small patches in the face

Today, fingerprints are still the most important modality

- From the image of a fingerprint the locations of ridge endings or ridge bifurcations are determined
- A set of about 30 of these locations is considered to be unique for an individual

Fuzzy cryptography

History of biometrics

Also, the iris is used and characterised by the texture of about 2000 bits

- Currently considered to be the best practical modality in terms of recognition performance

In the future, DNA might be utilised

- Probably in the form of so-called short tandem repeats (STRs)
- Would lead to a recognition error probability of 10^{-12} for unrelated individuals
- Not practical at the moment due to the lack of low-cost and quick DNA scanners
- Also, the use of DNA as biometric modality will rise privacy concerns

Fuzzy cryptography

History of biometrics

Generally, biometric techniques are extending from the forensic area to the civil area.

Instead of solving crime, they are utilised to prevent crime

Biometric systems will become omnipresent

- Biometric information of an individual will be stored in a large number of locations
- Usually not under the control of its owner

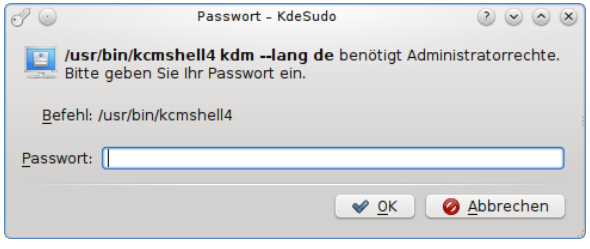
The ubiquitous storage of biometric information leads to a number of privacy risks such as identity theft

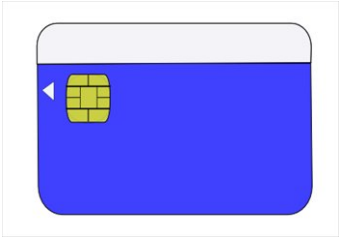
Outline

Use and implementation

Today, several devices and procedures are utilised to protect information

- Personal Identification Number (PIN)
- Passwords
- Smartcards
- Tokens to control access to services systems and information
- Subscriber Identity Module (SIM) cards
- Secure Socket Layer (SSL)
- ...





Use and implementation

It is becoming increasingly important to be able to authenticate physical objects and individuals

- Based on knowledge of a secret (PIN or password)
- Based on possession of a secret (smartcard)
- Based on biometric information (structure of a fingerprint, face, iris

This form of authentication requires a link to the physical world by measuring physical properties

These measurements are inherently noisy

- Temperature changes
- Humidity changes
- Light changes

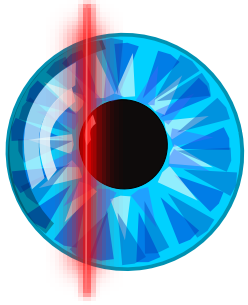
Use and implementation

During the last decades, a large range of security primitives has been developed.

An intrinsic property is, that these are extremely sensitive to small variations in their input.

However, in several security applications, noisy inputs can not be avoided.

The problem of using secret noisy data for security purposes is inadequately covered by traditional cryptographic primitives



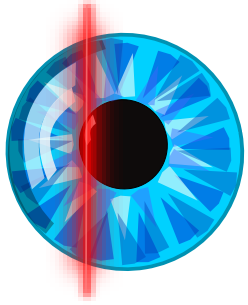
Use and implementation

In order to authenticate persons or objects, reference information is created

This information carries sensitive information about the individual from whom it is captured

Naive use and implementation of biometric systems might lead to severe privacy problems

For the physical object utilised, it is important that it does not leak information about secret properties of the object



Use and implementation

The three most important applications where noisy data is relevant for cryptographic purposes are

- Private biometrics
- Secure Key Generation
- Physical Unclonable functions

Use and implementation

Private Biometrics

Before an individual can use a biometric system, biometric reference information needs to be stored

During authentication, a live measurement is compared with the stored information

Storing biometric reference information in an unprotected manner will lead to security and privacy risks

Countermeasures are

- Not storing biometric data but construct it again each time the protocol is inferred
- Encrypt biometric information
- Transform noisy biometric reference information into a noiseless characteristic representation

Use and implementation

Secure key generation

Generate secret key from data transmitted over public communication channels and without sharing any secret in advance

Classical cryptographic setting: Attacker can eavesdrop without error

(introduced by Shannon)

Generalised: Attacker obtains noisy observation of messages

It is possible that the legitimate parties create a secure key while only transmitting public data

Use and implementation

Physical unclonable functions are inherently unclonable physical objects

PUFs map challenges to responses

- **Challenge** A stimulus applied to the PUF
- **Response** Reaction of the PUF obtained through noisy measurements

By embedding PUFs into devices, the devices become unclonable.

Challenge-response behaviour changes drastically when damaged

Instead of storing keys in the memory of a device, a key can be extracted from a PUF embedded in the device at the time required

Key is discarded when no longer required to minimise the time when it is vulnerable to physical attacks

Fuzzy cryptography

Physical unclonable functions

Definition

The object can be subjected to a large number of different challenges that yield an unpredictable response

The object is very hard to clone

Mathematical modelling of the challenge-response scheme is very difficult

It is hard to characterise the physical structure of the object

Fuzzy cryptography

Examples of PUFs

Silicon PUF

- During IC manufacturing there are always small variations even between ICs of the same wafer
- The variations do not harm the proper operation of the ICs
- They can be used as a source of randomness
- The challenge is a selection of a certain path on an IC
- The response is the delay time of a signal travelling along this path

Fuzzy cryptography

Examples of PUFs

Although it can not be considered as a proper object, the RF channel is in several properties very similar to PUFs

- It is able to produce a large number of unpredictable responses
  (provided that the attacker is not able to cross a critical minimum distance)
- It is hardly possible to clone it
- Mathematical modelling of the channel response is very difficult
  (provided that the attacker is not able to cross a critical minimum distance)
- Also, it is hard to describe the structure of a specific RF-channel
  (provided that the attacker is not able to cross a critical minimum distance)

Outline

Fuzzy cryptography

Utilise noise to improve security

Virtually all presently used cryptosystems can theoretically be broken

- by an exhaustive key-search
- Probably, they might even be broken due to novel algorithms
- Or by progress in Computer engineering

By exploring the fact that certain communication channels are inherently noisy, we can achieve secure encryption against adversaries with unbounded computing power

Fuzzy cryptography

Utilise noise to improve security

The security of essentially al presently used cryptosystem is based on at least two assumptions:

1. The computing resources of the adversary are bounded
2. The computational problem of breaking the cryptosystem is computationally infeasible

Both assumption are essentially not proven

1. The model of computation might even be unclear (recently demonstrated by quantum computers which are believed to be more powerful than classical computers)
2. Yet, no lower bound for the hardness of meaningful computational problems

Fuzzy cryptography

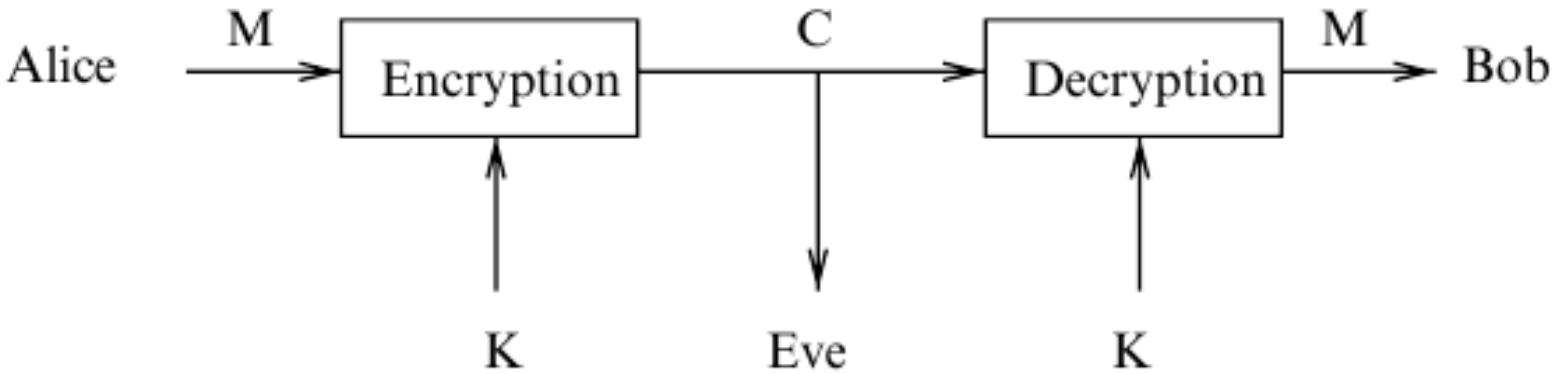
Utilise noise to improve security

Some unconditionally secure cryptosystems are proposed (secure against adversary with unbounded computing power)

Example: One-time pad

- Message $M = [m_1, m_2, \ldots, m_N]$
- Key $K = [k_1, k_2, \ldots, k_N]$ (uniformly distributed N-bit string)
- Cipher-text $C = [c_1, c_2, \ldots, c_N] = [m_1 \oplus k_1, \ldots, m_N \oplus k_N]$

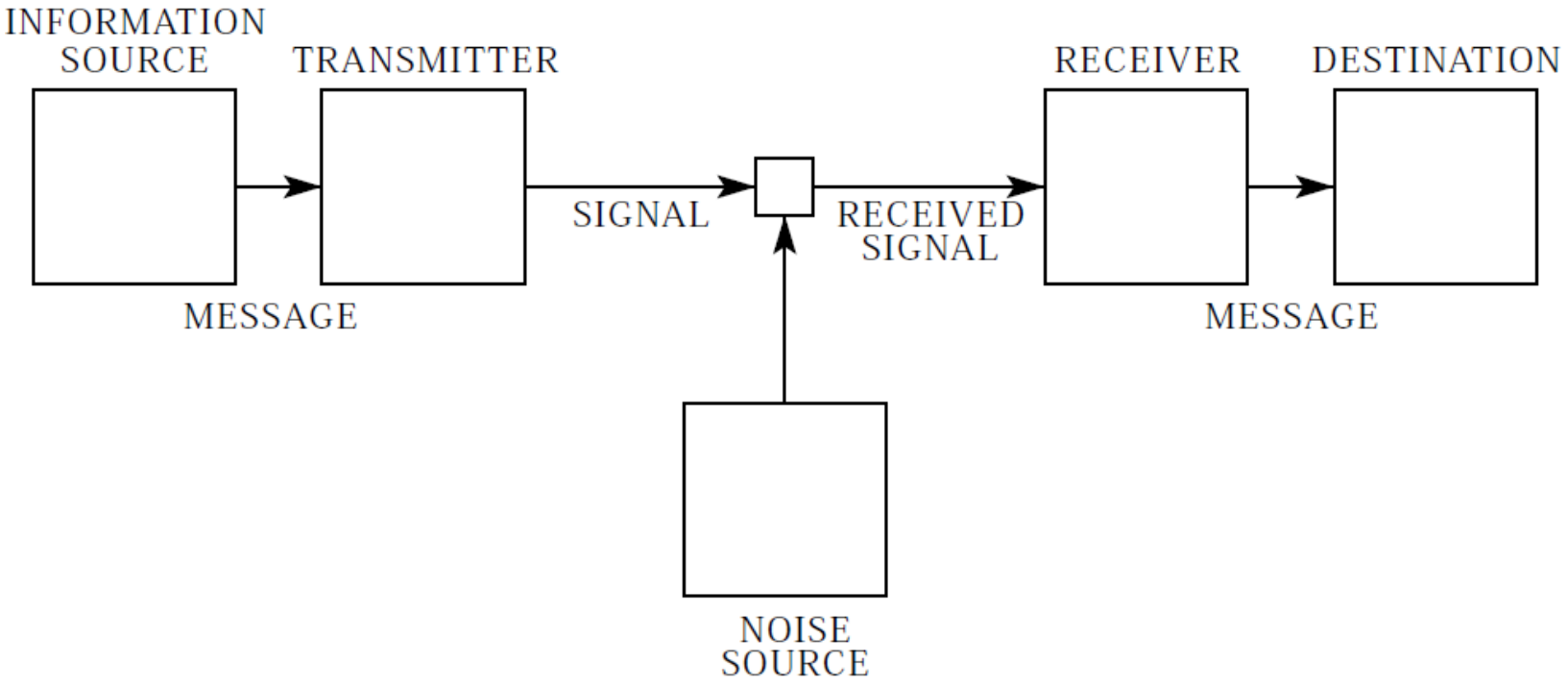
The one-time pad is perfectly secret



Fuzzy cryptography

Exkurs: Entropy

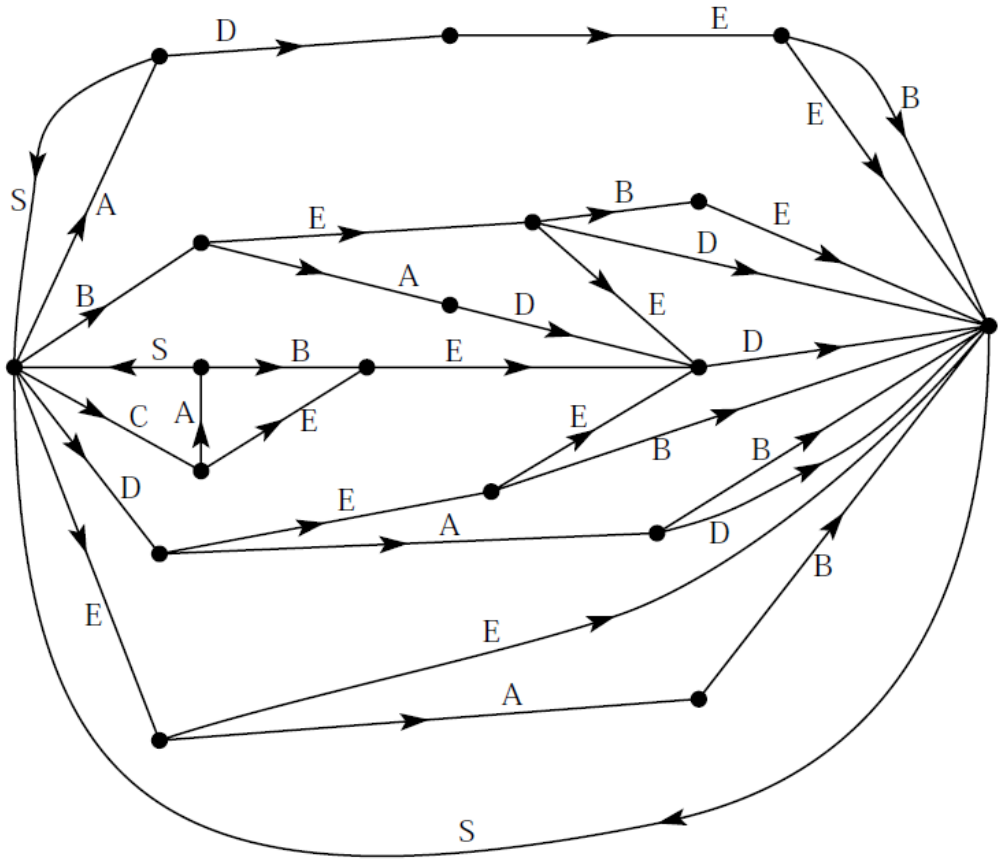
The fundamental problem of communication is to reproduce at one point a message created at another point



Fuzzy cryptography

Exkurs: Entropy – Discrete noiseless systems

A discrete information source can be represented as Markov process



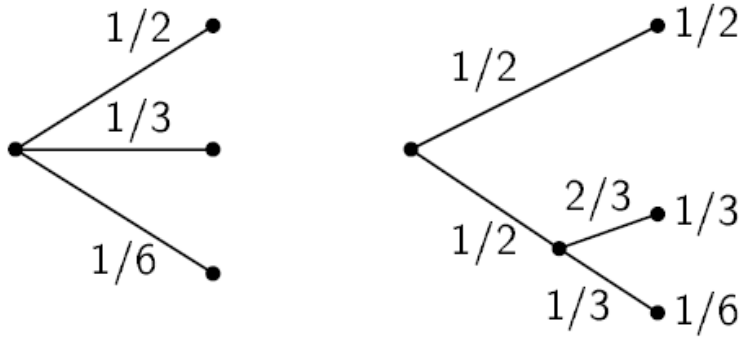
Fuzzy cryptography

Exkurs: Entropy – Discrete noiseless systems

For a measure $H(\cdot)$ of how much information is produced by an information source, assume a set of possible events with occurrence probabilities $p_1, \ldots, p_n$

For $H(p_1, \ldots, p_n)$ we require

1. H should be continuous in p_i
2. If all p_i are equal, H should be monotonic increasing function of n
3. If a choice is broken into two successive choices, the original H should be the weighted sum of the individual values of H



Fuzzy cryptography

Exkurs: Entropy – Discrete noiseless systems

- We can show that the only H satisfying all above assumptions is of the form[1]

$$H = -K \sum_{i=1}^{n} p_i \log_2 p_i$$

- We call this function the Entropy of a set of probabilities p_i

[1] C... A...athematical theory of communication, The Bell System Technical Journal, Vol. 27, 1948

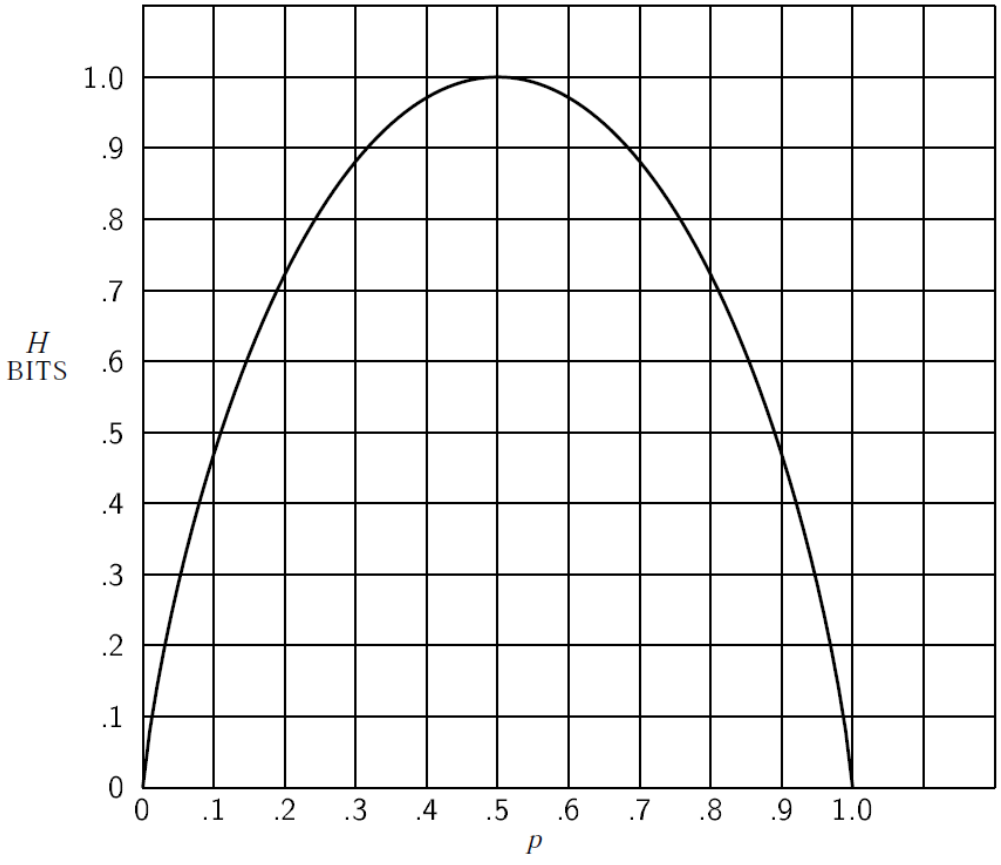
Fuzzy cryptography

Exkurs: Entropy – Discrete noiseless systems

Example:

- Entropy for the case p and $q = 1 - p$

$$H = -(p \log_2 p + q \log_2 q)$$



Fuzzy cryptography

Exkurs: Entropy – Discrete noiseless systems

$$H = -K \sum_{i=1}^{n} p_i \log_2 p_i$$

Properties of H

- $H = 0$ iff all but one p_i is 0
- maximum ($H = \log_2 n$) for $p_i = \frac{1}{n}$

Fuzzy cryptography

Exkurs: Entropy – Discrete noiseless systems

$$H = -K \sum_{i=1}^{n} p_i \log_2 p_i$$

Properties of H

- For the entropy of a joint event we have

$$\begin{aligned} H(x,y) &= -\sum_{i,j} p(i,j) \log_2 p(i,j) \\ H(x) &= -\sum_{i,j} p(i,j) \log_2 \sum_{j} p(i,j) \\ H(y) &= -\sum_{i,j} p(i,j) \log_2 \sum_{i} p(i,j) \end{aligned}$$

Fuzzy cryptography

Exkurs: Entropy – Discrete noiseless systems

$$H = -K \sum_{i=1}^{n} p_i \log_2 p_i$$

Properties of H

It is

$$H(x, y) \leq H(x) + H(y)$$

with equality only if the events are independent:

$$p(i, j) = p(i)p(j)$$

Fuzzy cryptography

Exkurs: Entropy – Discrete noiseless systems

$$H = -K \sum_{i=1}^{n} p_i \log_2 p_i$$

Properties of H

- The conditional entropy of y can be expressed as

$$H(y|x) = -\sum_{i,j} p(i,j) \log_2 p(j|i)$$

Fuzzy cryptography

Exkurs: Entropy – Discrete noiseless systems

$$H = -K \sum_{i=1}^{n} p_i \log_2 p_i$$

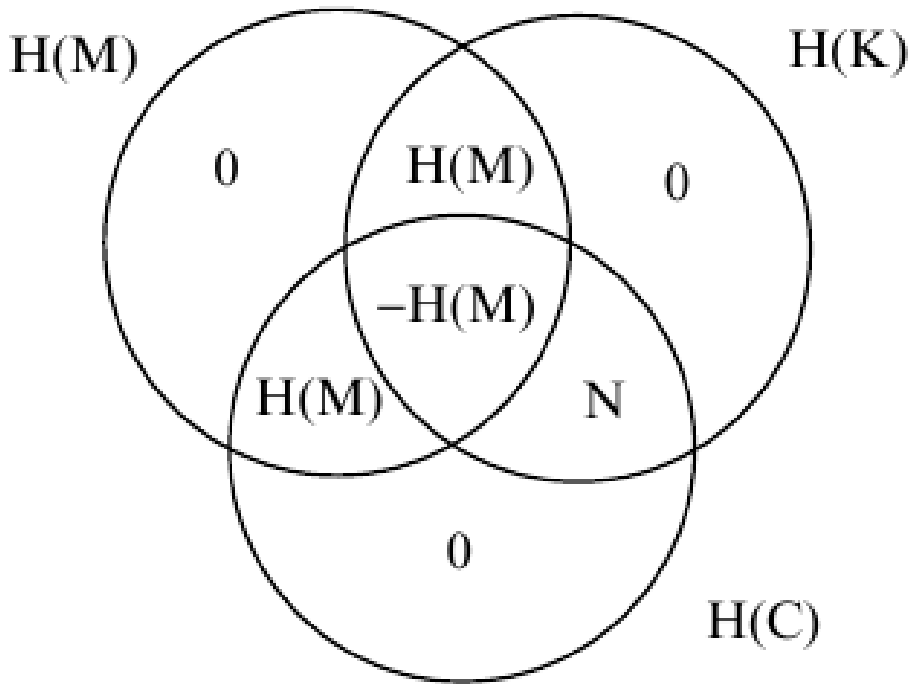
The entropy of a source is defined as the average of the conditional entropies weighted with the probability of occurrence of the states

$$\begin{aligned} H &= \sum_i P_i H_i \\ &= -\sum_{i,j} P_i p(j|i) \log_2 p(j|i) \end{aligned}$$

Fuzzy cryptography

Utilise noise to improve security

From an entropy diagram we can see that the one-time pad is perfectly secret



Fuzzy cryptography

Utilise noise to improve security

The price we have to pay for perfect secrecy is that

- communicating parties must share a secret key that is at least as long as the message
- and which can only be used once

The scheme is therefore quite impractical

However, Shannon showed that perfect secrecy can not be obtained in a less expensive way

The one-time pad is optimal with respect to key length

Fuzzy cryptography

Utilise noise to improve security

Consequently:

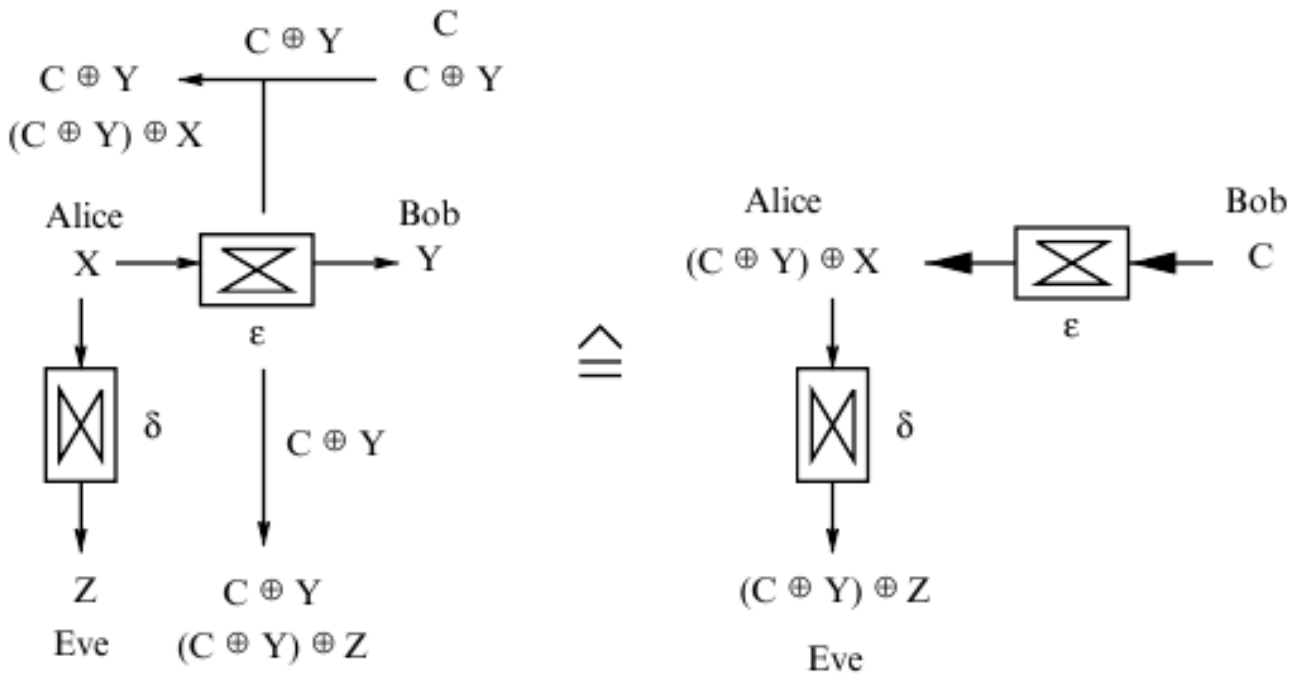
- Every perfectly secret cipher is necessarily as impractical as the one-time pad

However:

- The assumption that the adversary has perfect access to the cipher-text is unrealistic in general
- Every transmission of a signal over a physical channel is subject to noise
- We can utilise noise to achieve a perfectly secure communication at less cost

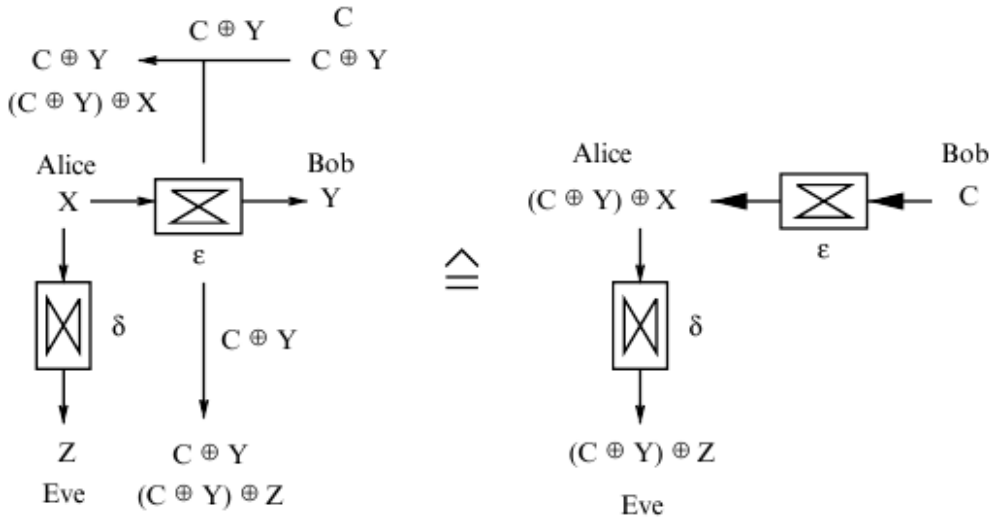
Fuzzy cryptography

Utilise noise to improve security



Fuzzy cryptography

Utilise noise to improve security



By inverting the direction of communication the noise in Eve's reception in increased above those in Alice's

Establishing of a secure key is possible over binary symmetric channel iff the noise in the reception of Eve's message is higher[2]

[2] ... wire-tap channel, Bell system Technical Journal, 54:1355-1387,1975

Outline

Fuzzy cryptography

Fuzzy Commitment

Traditional cryptographic systems rely on secret bit-strings for secure management of data.

When this key contains errors (e.g. due to noise or mistake), decryption will fail.

The rigid reliance on perfectly matching secret keys makes classical cryptographic systems less practicable in noisy systems.

Fuzzy commitment is a cryptographic primitive designed to handle independent random corruptions of the bits in a key.

Fuzzy cryptography

Fuzzy Commitment

Traditional cryptographic systems rely on secret bit-strings for secure management of data.

A cryptographic commitment scheme is a function

$$G : C \times X \to Y$$

To commit a value $\kappa \in C$ a witness $x \in X$ is chosen uniformly at random and $y = G(\kappa, x)$ is computed.

A decommitment function takes y and a witness to obtain the original κ

$$G^{-1} : Y \times X \to C$$

Fuzzy cryptography

Fuzzy Commitment

A well defined commitment scheme shall have two basic properties.

Binding It is infeasible to de-commit y under a pair (κ', x') such that $\kappa \neq \kappa'$

Hiding Given y alone, it is infeasible to compute κ

Fuzzy cryptography

Fuzzy Commitment

Fuzzy commitment is an encryption scheme that allows for the use of approximate witnesses

Given a commitment $y = G(\kappa, x)$, the system can recover κ from any witness x' that is close to but not necessarily equal to x.

Closeness in fuzzy commitment is measured by Hamming distance.

Fuzzy cryptography

Fuzzy Commitment

A fuzzy commitment scheme may be based on any (linear) error-correcting code

An error-correcting code consists of

Message space $M \subseteq F^a$ (F^i denotes all strings of length i from a finite set of symbols F)

Codeword space $C \subseteq F^b$ with $(b > a)$

Bijection $\theta : M \leftrightarrow C$

Decoding function $f : C' \rightarrow C \cup \perp$ (The symbol $\perp$ denotes the failure of f)

The function f maps an element in C' to its nearest codeword in C.

Fuzzy cryptography

Fuzzy Commitment

The noise introduced by a physical function may be viewed as the difference $c - c'$

The decoding function f is applied in an attempt to recover the originally transmitted codeword c

This is successful if c' is close to c. In this case we obtain $c = f(c')$

The minimum distance of the code is the smallest distance $d = Ham(c - c')$ between any two codewords $c, c' \in C$

Typically, it is possible to correct at least $\frac{d}{2}$ errors in a codeword

Fuzzy cryptography

Fuzzy Commitment

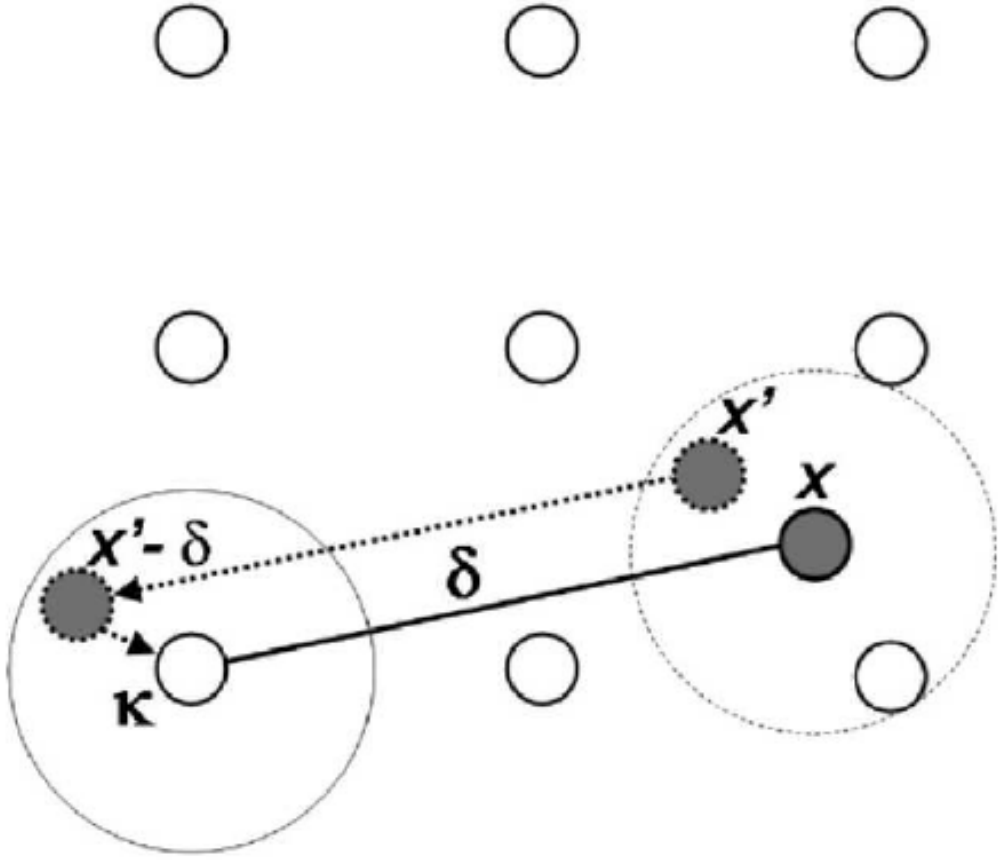
For fuzzy commitment, the secret key κ is chosen uniformly at random from the codeword space C. Then,

1. An offset $\delta = x - \kappa$ is computed
2. A one-way, collision-resistant hash function is applied to obtain $h(\kappa)$
3. $y = (\delta, h(\kappa))$ is made public
4. $\kappa' = f(x' - \delta)$ is computed
5. It is possible to de-commit y under a witness x' with $Ham(x, x') < \frac{d}{2}$

Once κ is recovered, its correctness my be verified by computing $z = h(\kappa)$

Fuzzy cryptography

Fuzzy Commitment



Outline

Questions?

Stephan Sigg
sigg@ibr.cs.tu-bs.de

Literature

- C.M. Bishop: Pattern recognition and machine learning, Springer, 2007.
- P. Tulys, B. Skoric, T. Kevenaar: Security with Noisy Data – On private biometrics, secure key storage and anti-counterfeiting, Springer, 2007.
- R.O. Duda, P.E. Hart, D.G. Stork: Pattern Classification, Wiley, 2001.