Information-Theoretic Cryptography

Ueli Maurer

ETH Zurich

BICI-INDAM 2005 International PhD School on Mathematical Aspects of Modern Cryptography, Sept. 4–9, 2005, Bertinoro.

Introduction

BICI-INDAM 2005 International PhD School on Mathematical Aspects of Modern Cryptography, Sept. 4–9, 2005, Bertinoro.

Information Theory in Cryptography

Information theory $\overset{?}{\subseteq}$ Cryptography

Information theory $\overset{?}{\supseteq}$ Cryptography

Information theory $\leftrightarrow$ Cryptography

Information theory vs. cryptography

Common features:

- Main math. tools: probability theory, algebra
- Crucial applications
- Fascinating science
- Fundamental concept of reductions

Distinguishing features:

- Average-case vs. worst-case analysis ($\forall$ adversaries)
- Computational hardness, complexity theory
- Verifyability of applications
- Viability of ad-hoc solutions
- Scientific communities

A classical prejudice

- Shannon proved that information-theoretic secrecy requires a (one-time) key at least as long as the message to be encrypted.
- This is completely impractical; hence we must resort to computational security.
- Computational security is ugly:
  - model of computation (e.g. Turing machine)
  - complicated definitions (e.g. polynomial time)
  - no ultimate proofs
- The main purpose of IT in cryptography is to prove impossibility results.

However, ...

- IT can also prove constructive (possibility) results for unconditional security.
- Many complexity-theoretic results are information-theoretic in nature
- ... if one interprets IT in a general sense.

  Often, Shannon entropy is not the relevant measure.

Assumptions in cryptography

- Every security proof is relative to assumptions!
  - Randomness exists (generation of secret keys)
  - Independence exist ($\nexists$ telepathy)
  - Computational intractability assumptions
  - Adversary's computing power and/or memory
  - Adversary's obtainable side information
  - Correct behavior (trustworthiness) of entities
  - Quantum theory is correct
  - Tamper-resistance of devices
  - Noise in communication systems
- Assumptions should be made explicit !
- Assumptions should be as weak as possible !

Goal in cryptography:

Information Theory Basics

Definition of entropy

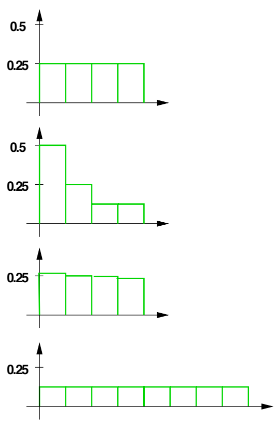
Entropy of a random variable X:

$$H(X) = -\sum_{x\in\mathcal{X}:P_X(x)\neq 0} P_X(x)\log_2 P_X(x)$$

Alternative notation: $H(X) = E[-\log P_X(X)]$

Theorem: $0 \leq H(X) \leq \log_2|\mathcal{X}|$

Entropy: some examples



entropy

2 bits

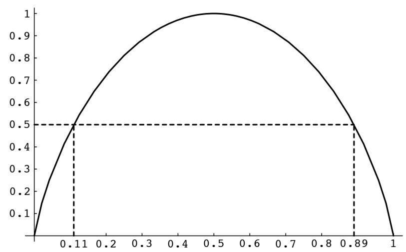
1.75 bits

1.99 bits

3 bits

Binary entropy function

$$h(p) = -p\log p - (1-p)\log(1-p)$$

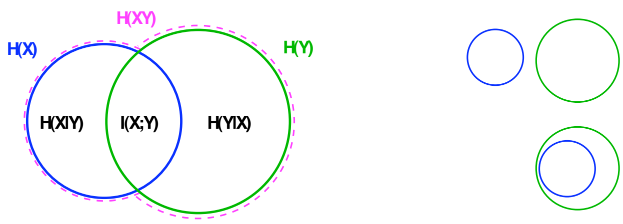


Conditional entropy and mutual information

H(X|Y) = H(XY) - H(Y)

I(X;Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(XY) = I(Y;X)

Theorem: $0 \leq$ H(X|Y) $\leq$ H(X) $\iff$ I(X;Y) ≥ 0

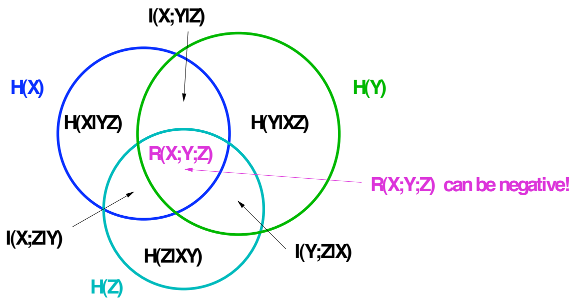


3 random variables: H(X|YZ) = H(XYZ) - H(YZ)

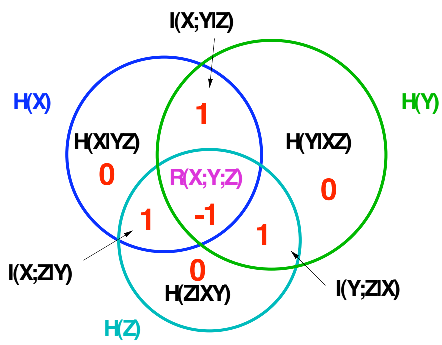
I(X;Y|Z) = H(X|Z) - H(X|YZ)

= H(XZ) + H(YZ) - H(Z) - H(XYZ)

Chain rule: H(XYZ) = H(X) + H(Y|X) + H(Z|XY)



Example: X,Y indep. random bits, Z = X $\oplus$ Y



Significance of Shannon entropy

Data Compression Theorem: Optimal data compression can compress the output of an information source arbitrarily close to its entropy. Error-free compression to below the entropy is impossible.

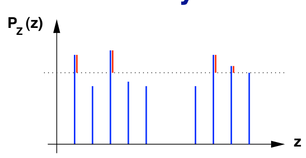
Example: An asymmetric binary source with $P(X_i = 1) = 0.11$ can be compressed to by a factor 2 because $h(0.11) = 0.5$.

Channel Coding Theorem: Optimal coding for a noisy communication channel allows to transmit information reliably at any rate arbitrarily close to the channel capacity

$$C = \max_{P_{\text{Input}}} I(\text{Input}\,;\,\text{Output})$$

Reliable transmission above capacity is impossible.

Distance from uniformity



$$\mathbf{d}(Z) := \frac{1}{2}\sum_{z\in\mathcal{Z}}\left|P_Z(z) - \frac{1}{|\mathcal{Z}|}\right|$$ (= sum of red quantities)

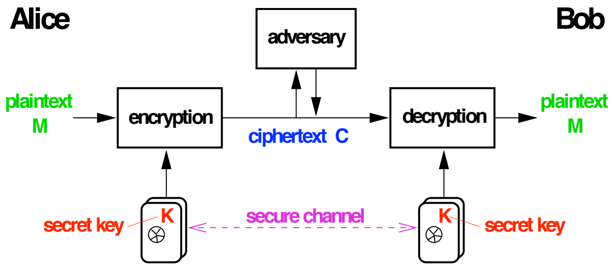
$$\mathbf{d}(Z|W) := E_W\left[\mathbf{d}(P_{Z|W}(\cdot|W))\right]$$

Lemma: One can define a uniform random variable Z' that is independent of W and such that Z $=$ Z' holds with probability $1 - \mathbf{d}$(Z|W).

Information-Theoretic Encryption and Key Agreement

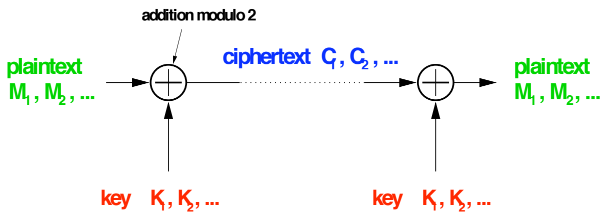
BICI-INDAM 2005 International PhD School on Mathematical Aspects of Modern Cryptography, Sept. 4–9, 2005, Bertinoro.

Symmetric cryptosystem

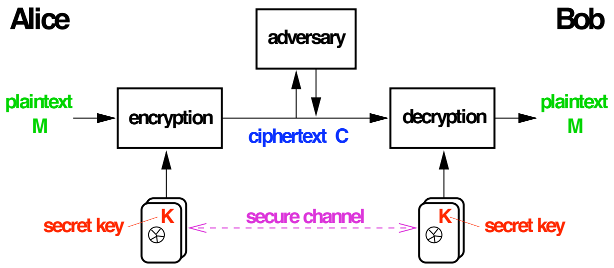


Perfect secrecy: I(M;C) = 0

One-time pad



Symmetric cryptosystem



Perfect secrecy: I(M;C) = 0

Theorem (Shannon): Perfect secrecy $\Rightarrow$ H(K) $\geq$ H(M)

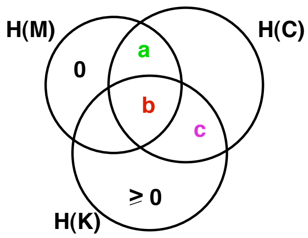
How to measure deviations from perfect? I(M;C)?

How to define computational security?

Shannon's theorem

Theorem: H(K) $\geq$ H(M) for every perfect cipher.

Proof:

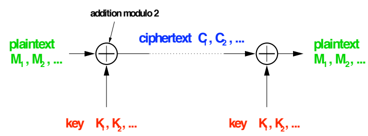


Decryptability: H(M|CK)=0

I(M;C) = 0 $\Rightarrow$ b = − a

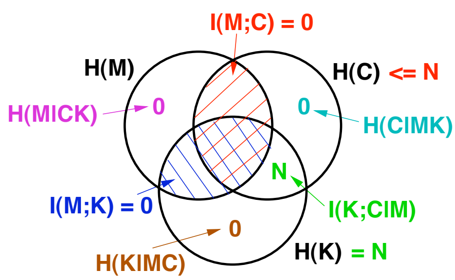
I(C;K) $\geq$ 0 $\Rightarrow$ c $\geq$ − b = a

H(K) $\geq$ H(M) by inspection



Theorem: The OTP is a perfect cipher for every P_M.

Proof:

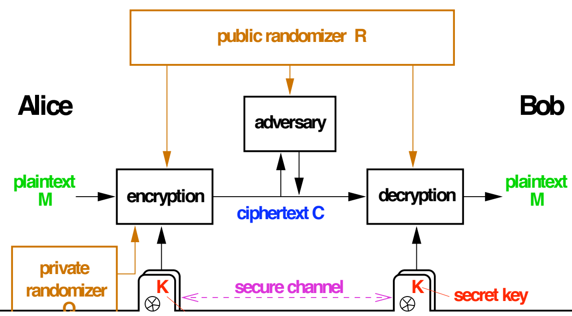


A discussion of Shannon's theorem

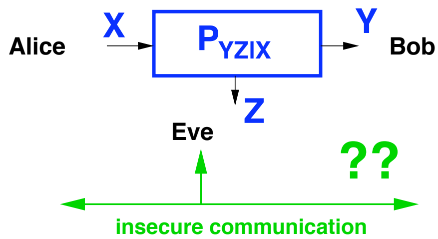
Significance of impossibility results:

- Assumptions should be general.
- No obvious modifications invalidating the impossibility result.
  - Randomization should be allowed!
  - Interaction (insecure) should be allowed!
  - Noise should be taken into account!

Symmetric cryptosystem with randomization



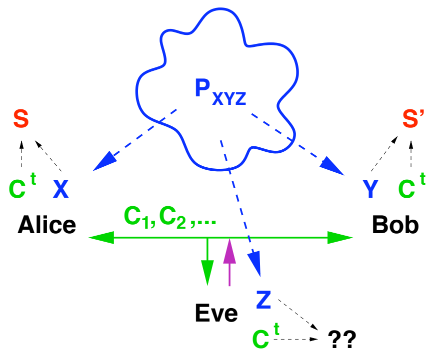
Wire-tap channels (Wyner, Csiszár-Körner)



Secrecy capacity $\geq$ I(X;Y) – I(X;Z)

It is 0 if Eve's channel better than Bob's

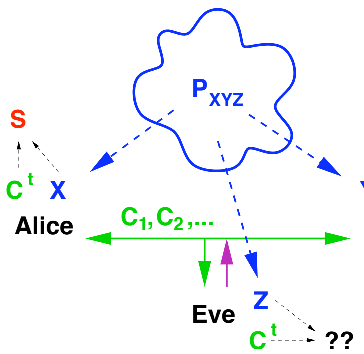
Secret key agreement by public discussion



S = S'

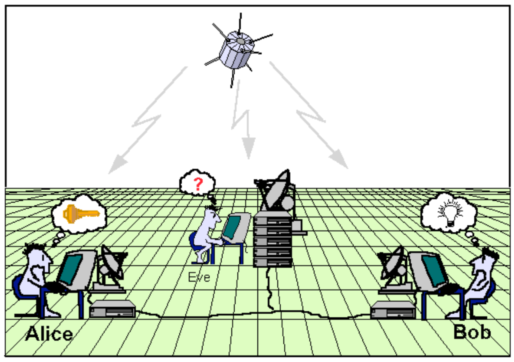
$I(S;C^t) = 0$

Secret key agreement by public discussion

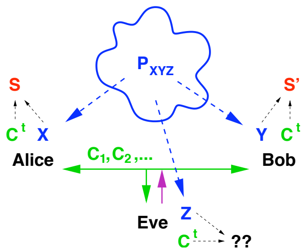


S = S'

$I(S;C^tZ)=0$



Secret key agreement by public discussion



S = S'

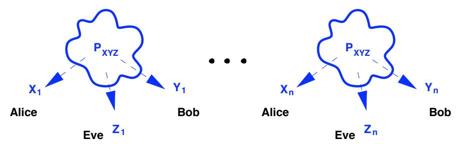
$I(S;C^tZ) = 0$

Theorem: $H(S) \leq \min [I(X;Y), I(X;Y|Z)]$

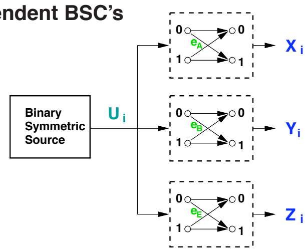
Corollary: The bound $H(K) \geq H(M)$ also holds in an interactive settings.

Corollary: A public-key cryptosystem cannot be information-theoretically secure.

Independent repetitions



Example: independent BSC's

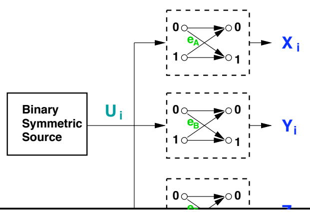


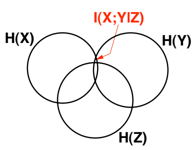
Secret-key rate

Definition: The secret-key rate of P_{XYZ}, denoted S(X;Y||Z), is the maximum rate at which A and B can agree on a secret key S.

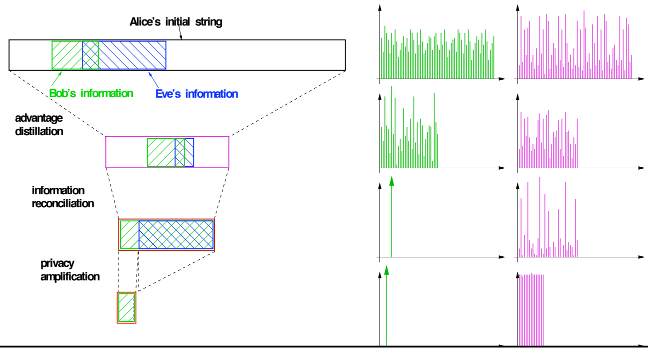
Theorem: S(X;Y||Z) $\geq$ max [0, I(Y;X) – I(Z;X), I(X;Y) – I(Z;Y)]

S(X;Y||Z) $\leq$ min [I(X;Y), I(X;Y|Z)]

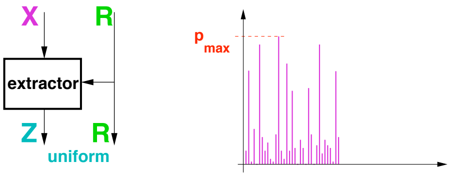




The three phases of secret key agreement



Privacy amplification



Goal: Generate uniform randomness

Deterministic: Only for some classes of P_X

Randomized, using a uniform catalyzer R:
$H_{min}(X) := -\log_2 p_{max}$ bits can be extracted with d(ZR) exponentially small.

R can be public.

Measuring deviation from perfectness

Question: Which is the right measure of deviation from perfect?

Proposal 1: I(M;C)

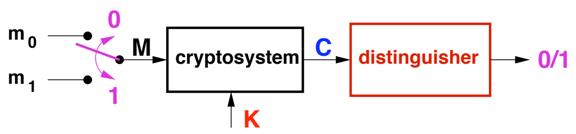
Proposal 2: Minimum of 1-P($\mathcal{E}$) such that I(M;C|$\mathcal{E}$)=0, maximized over message distributions P_M:

$$\max_{P_M} \min_{P_{M'C'}: I(M';C')=0} dist(P_{MC}, P_{M'C'})$$

Proposal 3: Maximal advantage, for any pair (m_0,m_1) of messages, of distinguishing the encryptions of m_0 and m_1:

$$\max_{m_0,m_1} dist(P_{C|M=m0}, P_{C|M=m1})$$

Proposal 4: Simulatability definition.



Comptational security?

A cryptosystem is **indistinguishability secure** if

- for all messages m_0 and m_1,
- for any efficient distinguisher,

the advantage in distinguishing the encryptions of m_0 and m_1 is negligible.

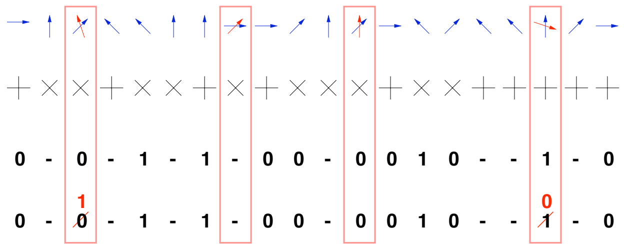
efficient = polynomial time

negligible = vanishes faster than inverse to any polynomial

Quantum Cryptography: A Glimpse

BICI-INDAM 2005 International PhD School on Mathematical Aspects of Modern Cryptography, Sept. 4–9, 2005, Bertinoro.

Quantum cryptography: example



first row: photons sent by Alice
second row: bases selected by Bob
third row: bits generated by Alice
fourth row: bits generated by Bob

Quantum cryptography: some explanations

- Alice and Bob are connected by a conventional insecure but authenticated communication channel as well as an optical fiber allowing Alice to send photons to Bob. Eve has access to the fiber.
- The polarisation of a photon can encode information, but due to the laws of quantum physics, only two states can reliably be distinguished by any measurement. Hence one can transmit reliably only 1 bit of information by encoding the two bits in orthogonal polarisations.
- Two different bases for sending a bit are defined: the horizontal/vertical basis and the diagonal ($45^o/135^o$) basis.
- Alice sends a sequence of random bits, each in a random basis. Eve cannot measure exactly which of the 4 states was transmitted.
- Bob measures each received photon in random basis and tells Alice which bases he has used. Alice announces for which bits Bob used the right basis and hence knows Alice's bits. Using error correction and privacy amplification, Alice and Bob can extract a secret key.
- One can prove that Eve has only a choice between performing too strong measurements and therefore being detected by Alice and Bob with high probability, or obtaining essentially no information about the derived key.