Information security: an Overview on Cryptography
MSc spring 2017

István Vassányi PhD
vassanyi@irt.vein.hu

(Lecture notes excerpts for 2 lectures)

The basics of modern cryptography

- Aspects of data security: organizational, physical, IT. The goal is to establish a system that is uniformly safe in all aspects. The principle of design is that the time or cost of a successful attack be more than the value of the information.
- Classic **mono-alphabetic systems**, Caesar, S-box, attacks via symbol statistics and improvements in the Middle Ages
- The dangers of **leaning over a decayed view-tower railing**: the case of Scottish Queen Mary, 1586
- The basic scheme of cryptography: role of the **plain text, ciphertext, key**, encryption, decryption, sender, channel, receiver, attacker
- The goal of the attacker is to find the key. Types of attacks: cíphertext, plain text, chosen plain text, chosen ciphertext
- The **brute force** ciphertext attack
- We consider an algorithmic attack successful if the number of possible keys could be decreased by 1
- Other attacks via manipulation of channel/network elements: **replay, cut and paste, man in the middle** (identity theft)
- **Perfect secrecy** against ciphertext attack
    - Defined by Shannon as $|X|=|Y|=|K|$ ie. the number of keys, cipher texts and plain texts must be equal
    - $P(K_i)=\frac{1}{|K|}$ ie. we choose the key for each message with a uniform probability
    - Why the brute force attack does not work: the attacker does not know if (s)he found the key

      cipher: PEFOGJ = PEFOGJ
      *key*1: PLMOEZ *key*2: MAAKTG
      plain: ATTACK ≠ DEFEND
      (both plain texts make sense)
    - **Key management problem** limits application to diplomacy (the Moscow–Washington hotline established after the 1962 Cuban missile crisis)

- **Secret key cryptosystems**

We do not use a perfect crypto system, rather, we force the attacker to try all the keys i.e. implement a full brute force attack AND we choose $|K|$ large enough to exclude a practical attack. This is called the principle of **computational secrecy**. We find the key size by the Thermodynamic Limitations. 256 bits needs more energy than that produced by the Sun in its whole life time.

Why we didn't need to try all the keys to break Caesar: the **language entropy**

$$H_L = \lim_{n\to\infty} \frac{H(A_1, A_2, \ldots A_n)}{n}$$

For English,

n	$\frac{H(A_1, A_2, \ldots A_n)}{n}$
1	4 bit
2	3.56 bit
3	3.30 bit

The language redundancy:

$$R_L = 1 - \frac{H_L}{\log |A|}$$

We make use of this redundancy to exclude several candidate keys at a time, by trying a key-**fragment** that produces an inacceptable plain text fragment. For a mono alphabetic cipher:

$$n_0 = \frac{\log |K|}{R_L \log |A|}$$

For English, $n_0 = 25$.

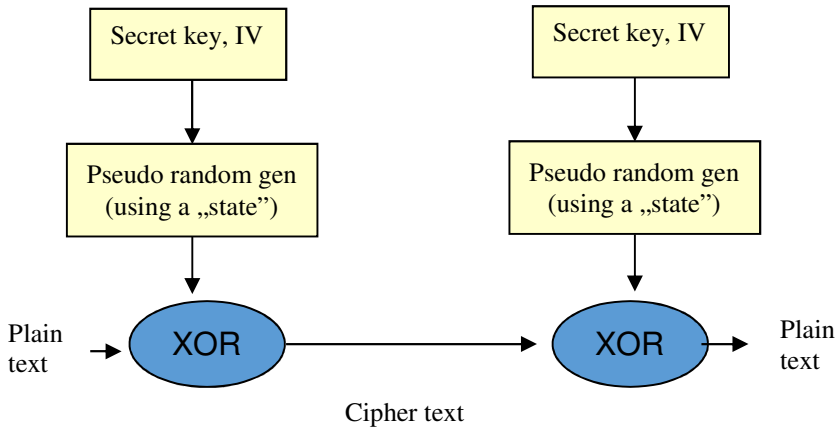
A **symmetrical crypto system** uses the same key at both sides and the encryption and decryption algorithm is virtually the same.

Block Ciphers

- **product cipher** proposed by Shannon: iterations on a kernel with P-box (linear), S-box (nonlinear), encryption (XOR) with an iteration key. Expected **avalanche effect**.
- The parts (bits) of the key, the cipher block, and the plain text block will be statistically independent, so only a full brute force attack is possible.
- You already learnt the details and various operational modes of AES, DES like ECB, CBC, OFB to prevent cut and paste attacks
- The weak point is the true random key. Sources of random keys. The **Golomb criteria** (A5/1 DEMO)
- Secret key extension, generating several session keys from a single secret key: **Key Encrypting Key**

Stream Ciphers

Faster than block ciphers. The principle is a **pseudo random generator** seeded by a secret key and an **initialization vector**.



A block cipher in OFB mode can always be transformed into a stream cipher.
We expect that

1. The state cannot be deduced from the pseudo random bit stream
2. The secret key cannot be deduced from the state
3. The bit stream should be random according to the Golomb criteria

- RC4: (1987) uses 1-byte blocks. The secret key is a permutation of the 0..255 series in an array, so |K| = 256! resulting in a key strength of 1680 bits. The period is very long. In each iteration, a permutation of the array is applied and a coding byte is selected and XOR-ed with the plain text byte. Vulnerabilities found in 2014, 2016.
- A5/1 and improved versions. The 128 bit secret key is stored on the SIM card. The session key is generated on a challenge-response basis and used to initialize the LFSR's.

Public key Ciphers

There is no safe channel between A and B, not even for key exchange. All public key methods are based on a **trap function** i.e. an operation that is easy to perform but very hard to invert.

- **Discrete logarithm problem** (DLP). If a, x, and p are known, then $y = a^x \bmod p$ is easy to compute, but a, y, and p given it is hard to find x for large numbers. The basis of the Diffie-Hellman key exchange protocol and the Elgamal algorithm.
- **Prime factorization problem** (PFP). If p and q big primes, then it is easy to compute $n = p \cdot q$ but n given it is hard to find p and q. The basis of RSA.

Attacks against public key ciphers (PKC) are algorithmic instead of brute force, and in contrast to the Thermodynamic Limitations, it is hard to predict their future. PKC methods are therefore less secure than secret key systems, and also much slower. PKC is therefore applied only in combination with secret key systems, forming **a hybrid cryptosystem** like SSL. Never used for mass encryption. DEMO: SSL logfile

You already learnt RSA. Attacks against RSA:

- Exploiting arithmetic errors like a too small n
- Chosen ciphertext attack exploiting a protocol error. (**Oracle service**, the sea battle at Midway 1942)
- Timing attack: by measuring computation time, the attacker could exclude a large number of possible keys

Diffie-Hellman key exchange

1. Alice and Bob choose a big prime p publicly and a small g generator of the cyclic group based on p:

$$\forall x \in GF[p]\ \exists\ k\colon g^k = x\ mod\ p$$

2. A chooses a random $a < p$ and sends to B $k_a = g^a mod\ p$ számot. Similarly, B sends $k_b = g^b mod\ p$ to A publicly. The attacker must solve the DLP.
3. Both A and B compute $K = {k_b}^a\ mod\ p = {k_a}^b\ mod\ p = g^{ab}\ mod\ p$ and use it as the secret session key for the communication.

ElGamal crypto system, 1985, a modification of D-H

- Method on ppt
- **demo**: ElGamal encryption (mathematica)
- also used in DSS (Digital Signature Standard) for signature

Signatures and certificates in a PK Infrastructure

- The principle of **digital signatures** using PKC: the private and the secret key are unique to each other. The signer encrypts only a **hash**, and not the whole document for efficiency. The result is **non-repudiation** and **integrity** of the message. The digital signature in Hungary is legally valid as of Act XXXV/2001. Problems: private keys are stolen, certificates expire (see later). Paper based signature still prevails.
- the **man-in-the-middle-attack** against PKC
- the use of **certificates** to prevent the man-in-the-middle-attack. The **Certificate Authority**. Parts by the X509 standard: CA related data, expiration, name and other particulars (data) of owner, method and value of owner's public key, algorithm used for signing, plus a signature prepared by the CA for all these data as a single document
- certificate authentication systems: hierarchical (**chain of trust** followed up to a **trust anchor**), or Pretty Good Privacy, PGP in a **web of trust**
- certificates are to be checked before use at a trusted Online Certificate Status Protocol (OCSP) server for the **certificate revocation list**
- recommended for personal privacy: dedicated hardware generating, storing and using the secret key, also storing the certificates of others.

Message digests

MD algorithms ae fast (see table below).

You learnt SHA 1-2 already.
Attacks:

- Digest -> message: **preimage** attack)
- Message -> another message with the same digest : **second preimage** attack
- Two messages with the same digest : **collision** attack
- The **birthday** attack
- **Rainbow tables** (*K14.4*) for breaking max. 8 character password hashes. The table is specific for a hash, password length, and character set. Prevention: **salting** or multiple application of the hash (**key stretching**).

Relative speed of cryptosystems

method	speed	assessment
Message digest (hash)	3	fastest
Stream cipher	2	faster
Block cipher	1	fast
Public key cipher	0.02	Very slow

Recommended key sizes

Key size for secret key CS	*Security level*	*RSA key size*	*ECC key size*
72	Cracked in a short time by simple methods	1008	144
80	"theoretically" sufficient	1248	160
96	Recommended "bare minimum""	1776	192
112	Recommended "sufficient minimum"	2432	224
128	Sufficient except top secret documents	3248	256
256	Sufficient even for top secret documents	15424	512

Recommended elements of a modern commercial system:

Algorithm	*Role*
RSA 3072-bit or larger	Key Establishment, Digital Signature
Diffie-Hellman (DH) 3072-bit or larger	Key Establishment
ECDH with NIST P-384	Key Establishment
ECDSA with NIST P-384	Digital Signature
SHA-384	Integrity
AES-256	Confidentiality

It is not always possible to provide the sufficiently large key sizes for RSA due to resource limitations e.g. in smart card applications. This gives rise (or should give rise…) to the ECC.

Elliptic Curve Cryptography

The ECDLP replaces the DLP.
(see separate slides)

New Stream Ciphers

- Rabbit
- Grain

(see separate slides)

"Quantum" Cryptography

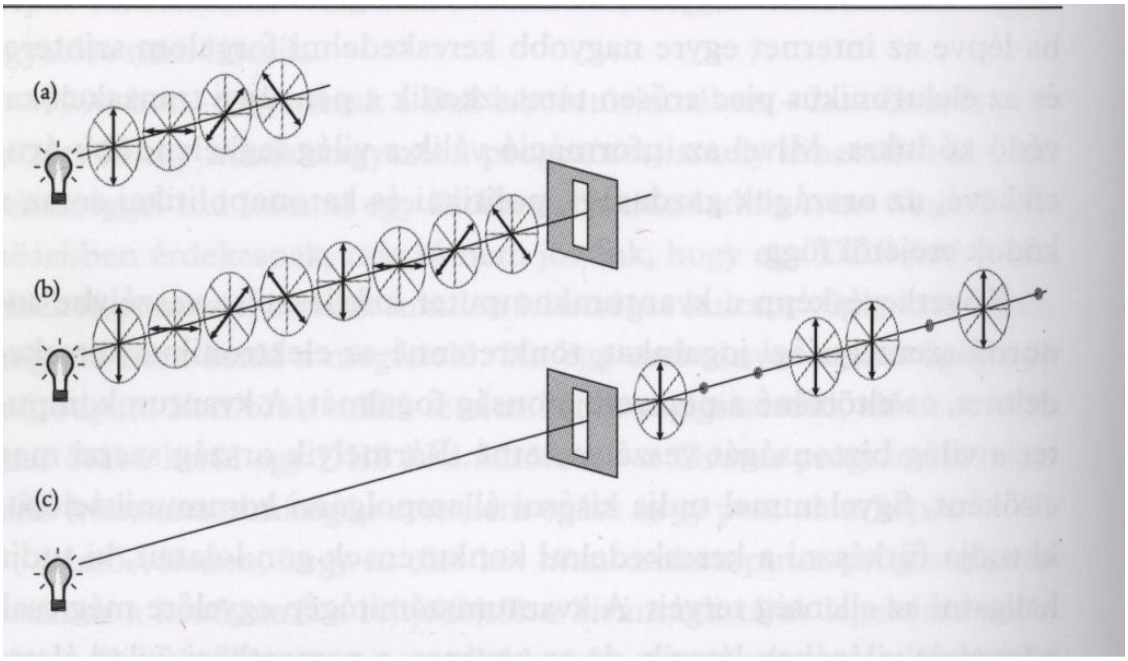
1. Quantum Key Distribution Networks, QKDN. The BB84 protocol for key exchange.
2. The Quantum Computer for PFP

Quantum Key Distribution Networks

A polarized photon will pass a diagonal filter with a probability of ½.

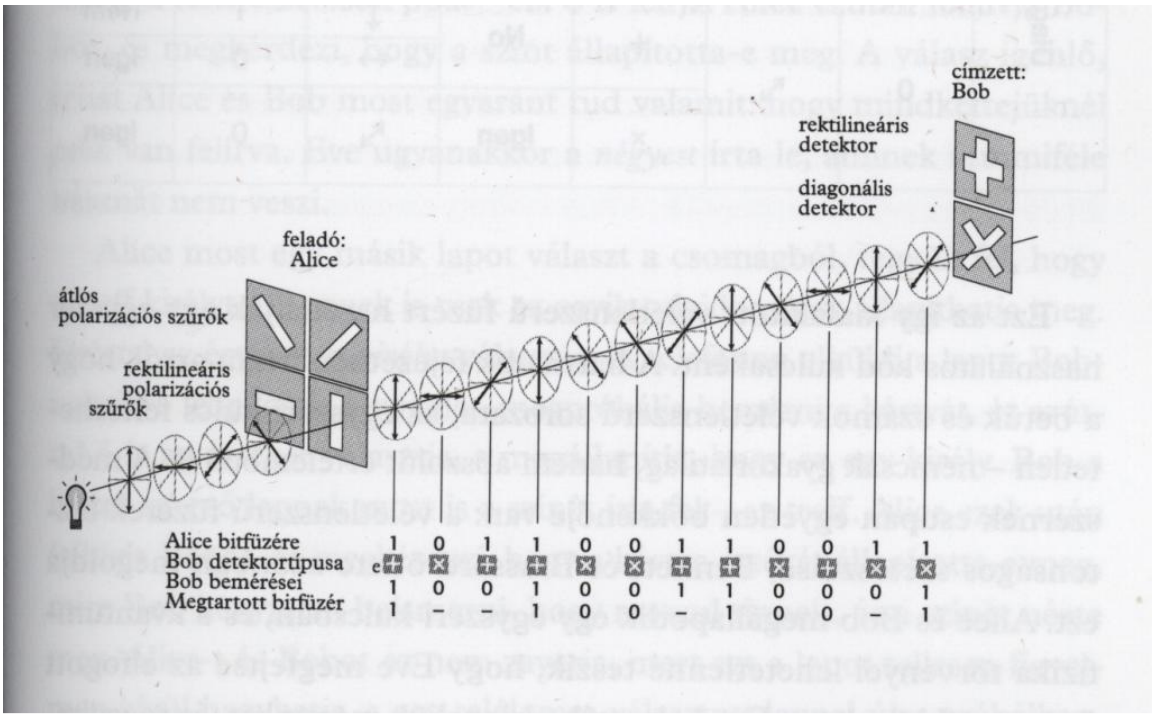
As of Quantum Theory, "The measurement always modifies the state of the system".

"no-cloning" theory



BB84 protocol steps:

0. Authentication (via certificates)
1. Both parties choose a random filter set (basis), and Alice uses the basis to produce a photon from each bit of the key, Bob uses his basis to detect the photon



Alice sémája	Alice bitjei	Alice ezt küldi	Bob detektora	Megfelelő detektor?	Bob megállapítása	Bob bitje	Helyes Bob bitje?
Rektilineáris	1	↕	+	Igen	↕	1	Igen
			×	Nem	⤢	1	Igen
					⤡	0	Nem
	0	↔	+	Igen	↔	0	Igen
			×	Nem	⤢	1	Nem
					⤡	0	Igen
Diagonális	1	⤢	+	Nem	↕	1	Igen
					↔	0	Nem
			×	Igen	⤢	1	Igen
	0	⤡	+	No	↕	1	Nem
					↔	0	Igen
			×	Igen	⤡	0	Igen

2. discuss basis choices publicly and keep only bits at matching positions
3. in order to detect eavesdropping, publicly discuss a small set of key bits. Eve will spoil the key bit in ¼ of the cases.

Alice's random bit	0	1	1	0	1	0	0	1
Alice's random sending basis	+	+	×	+	×	×	×	+
Photon polarization Alice sends	↑	→	↘	↑	↘	↗	↗	→
Eve's random measuring basis	+	×	+	+	×	+	×	+
Polarization Eve measures and sends	↑	↗	→	↑	↘	→	↗	→
Bob's random measuring basis	+	×	×	×	+	×	+	+
Photon polarization Bob measures	↑	↗	↗	↘	→	↗	↑	→
PUBLIC DISCUSSION OF BASIS								
Shared secret key	0		0			0		1
Errors in key	✓		✗			✓		✓

By discussing n bits, the probability of detecting the attack is $P = 1 - \left(\frac{3}{4}\right)^n$, $\longrightarrow 1$.

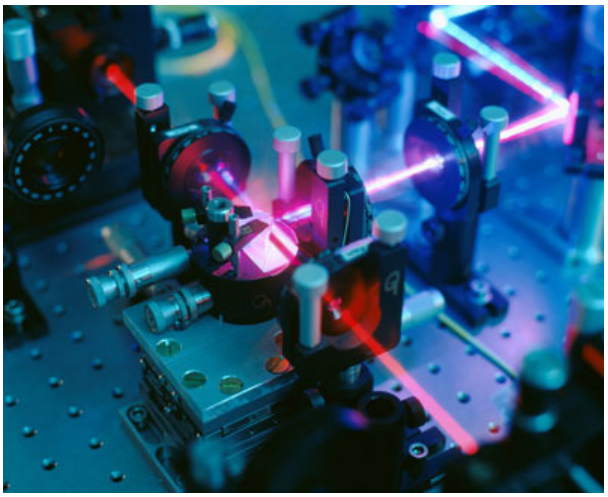
information reconciliation/ privacy amplification: bits can be shared even in the presence of the attacker

Attacks:

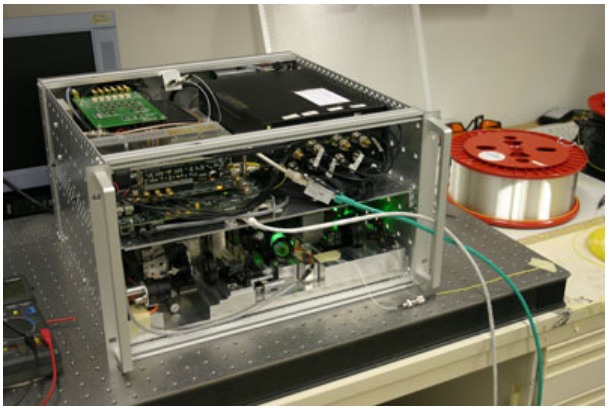
- Eavesdropping
- man-in-the-middle
- photon splitting
- hacking against the RNG and protocol
- DoS (cut the optical cable)

State of the art:

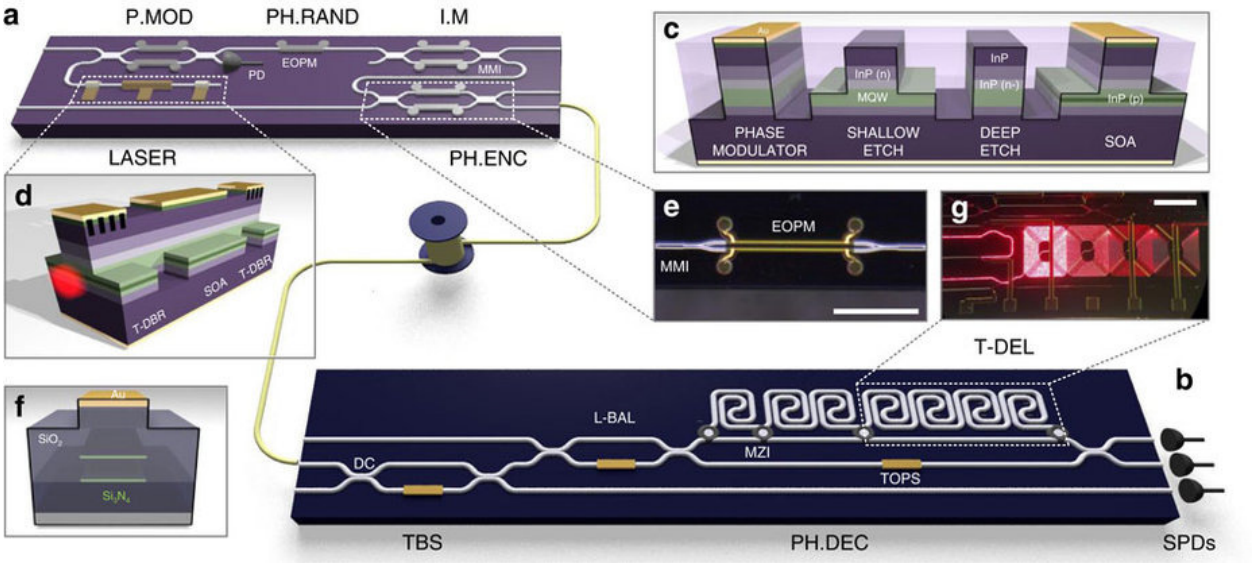
- estimated secret key rates up to 568 kbps, for an emulated 20 km fiber link.
- Several projects running in Japan and Europe



Photon source



QKD prototype



Chip-based quantum key distribution

The “Quantum Computer”

- Computation is performed by nature’s principle of minimum energy
- The problem is formulated in the form of a specially shaped magnetic field
- When the state of the qubits stabilizes, it can be queried via ‘entangled qubits’
- Expensive technology with superconducting niobium at 0.01 K
- Applied for PFP. In 2015, 56153=233*241, was factored
- Google, NASA