Chapter 2

Types of modern encryption

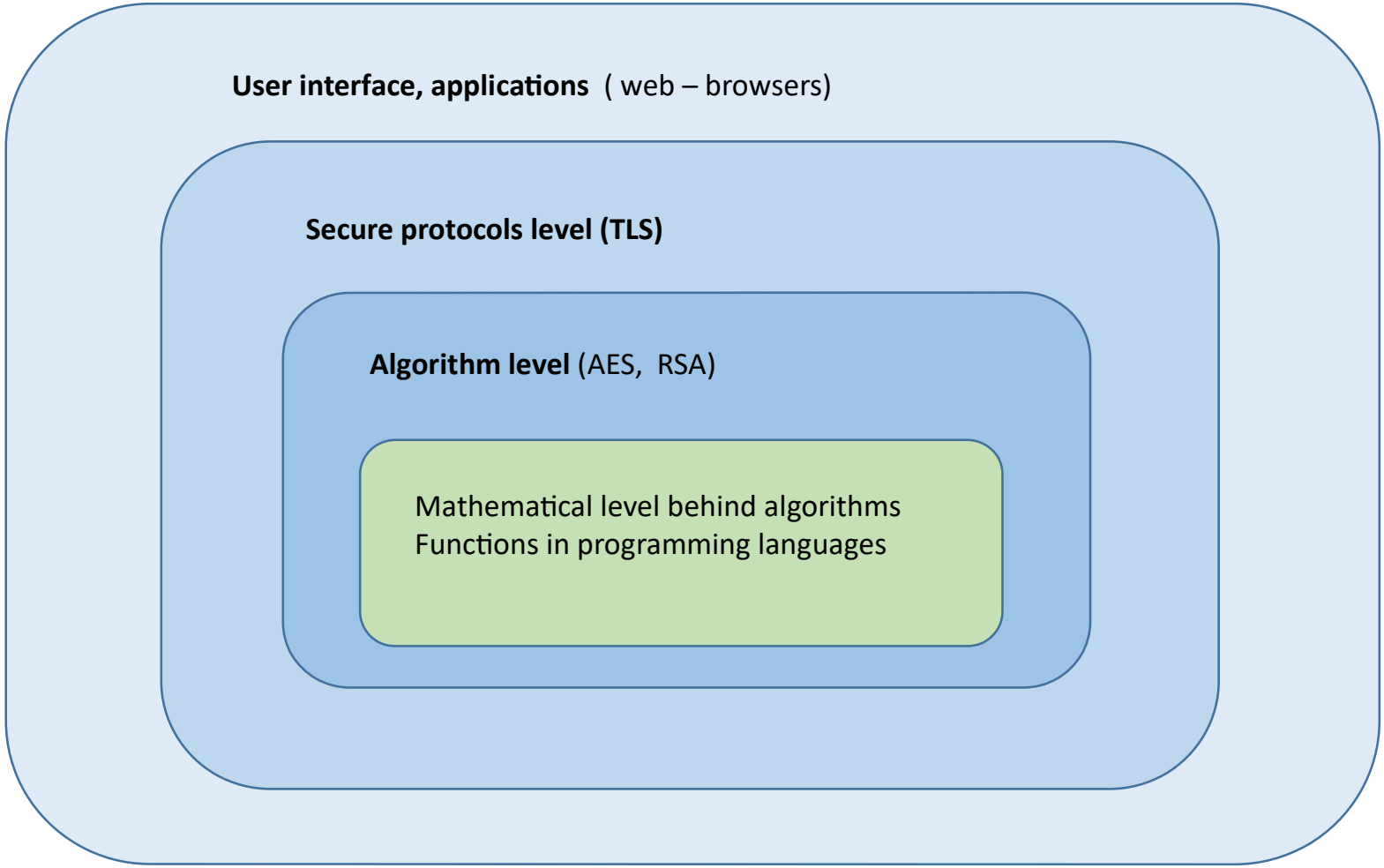
Secure connections seen from different levels

- User interface, applications
- Secure protocols level
- Algorithm level
- Level of mathematical structures

Block ciphers

Stream ciphers

Encryption at different levels



User level (browsers)

Users know that the connection is secure from the lock of URL row of the browser. If the lock is green, the connection is secure and encrypted and the server is authenticated. Red lock means possible security problems.

The encryption software which provides the secure connection is usually TLS.

Secure protocols level

In Mozilla Firefox browser clicking the lock reveals following information:

Tekniset tiedot
Yhteys salattu (TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384, 256-bittinen avain, TLS 1.2)

TLS ver 1.2 is a hybrid encryption software, which uses many algorithms for different functions. More than 95% of secure connections over Internet use TLS.

The basic functions of a hybrid encryption software are:

1. Authentication
2. Key exchange (agreeing of symmetric key)
3. Encryption of data transfer
4. Digital signatures

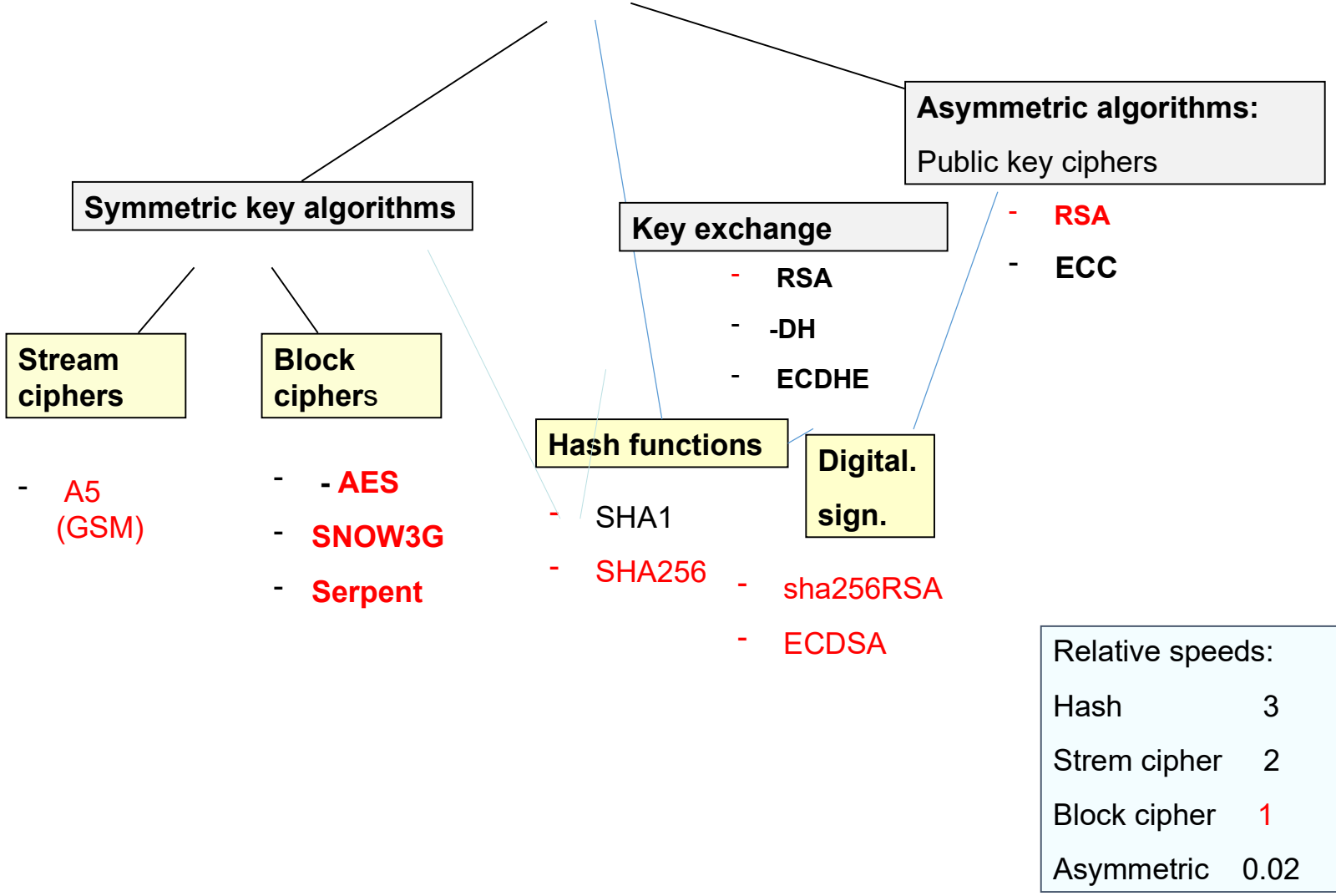
Algorithmic level

In this course one major focus is to learn about algorithms

Below list of algorithms is from Nordea Bank's TLS - connection

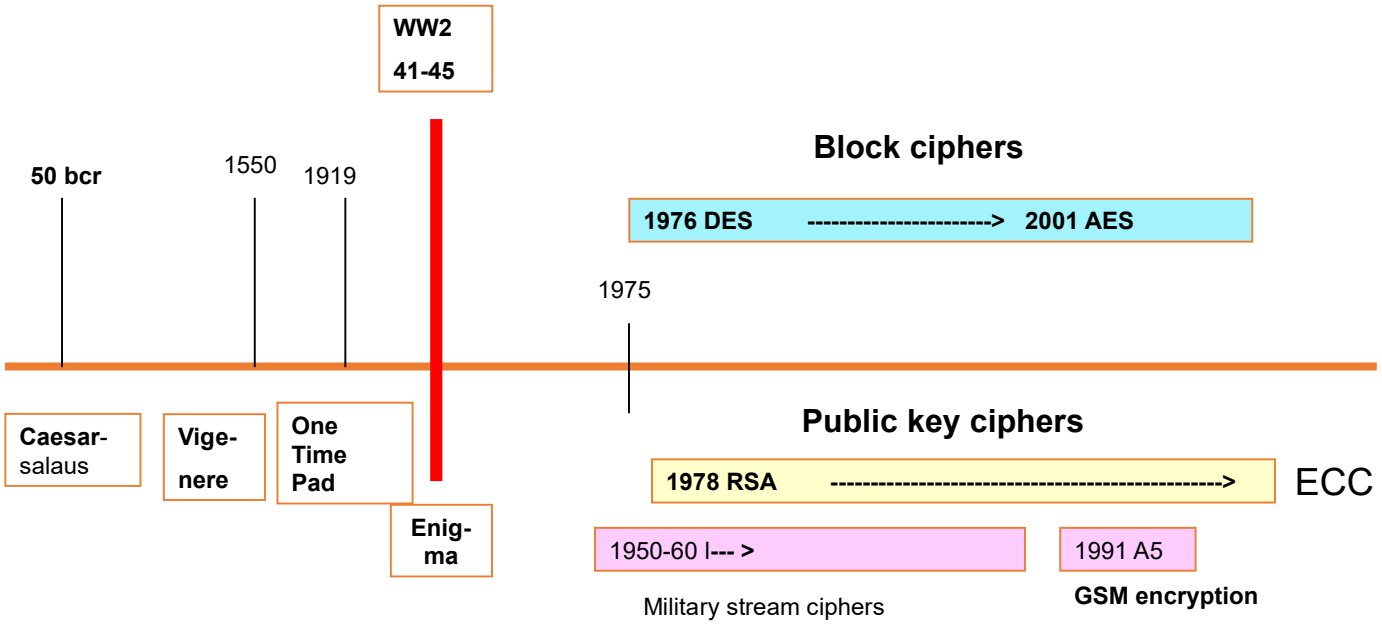
Function	Algorithm name
Server authentication	**RSA**
Key exchange	**ECDHE**
Encryption of transmitted data	**AES256 GCM-moodi**
Digital signature	**sha384RSA**

Algorithm types



Relative speeds:	
Hash	3
Strem cipher	2
Block cipher	1
Asymmetric	0.02

Time line of encryption methods



Algorithms have long lifetime. RSA has been standard for 45 years in public key cryptography. Block cipher standard has changes only once during 45 years. Reasons for conservatism: 1. Changing algorithms costs money. 2. The security of algorithms are based on mathematical proofs, they have been tested for decennies.

4. Mathematical level

The algorithms of TLS and other modern software require lots of mathematics

1. Fermat's and Euler's theorems, Extended Euclid's algorithm	
2. Random number generation	**pseurodandomness**
3. Prime generation, tests	**Fermat and Rabin- Miller tests**
4. Modular arithmetics	**number sets Z_n**
5. Fast exponentiation mod n	**Powermod - algorithm**
6. Inverses mod n	**Euclid's extended algorithm**
7. Cyclic group Z_p^*	**group theory**
8. Elliptic curves	**group theory**

Symmetric encryption algorithms

Classification

1. Block ciphers

-used both in wired and mobile communication
-Diffuusion and confusion principles
-Permutation-substitution networks
-Modes of operation
-AES : block cipher standard

2. Syncronic stream ciphers

- no longer widely used
- GSM (G2)– encryption A5

Block ciphers

- Central part of secure connections. Block ciphers encrypt transmitted data reliably and effectively
- Provides fast symmetric encryption in which message is encrypted in block (size usually 128 bits)
- The present block cipher standard is AES (since 2001). It is used both in secure wired connections and in mobile networks

History of block ciphers

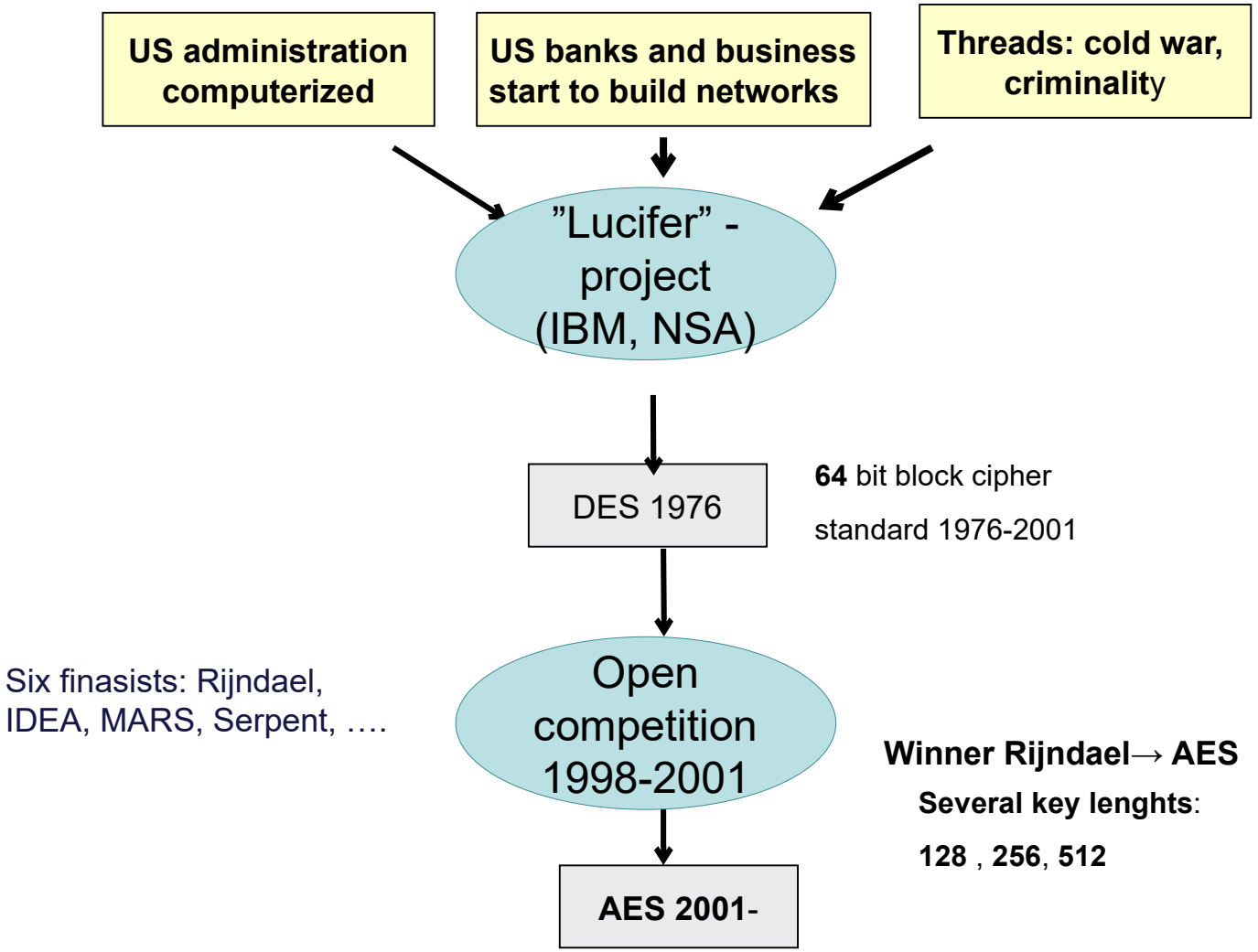
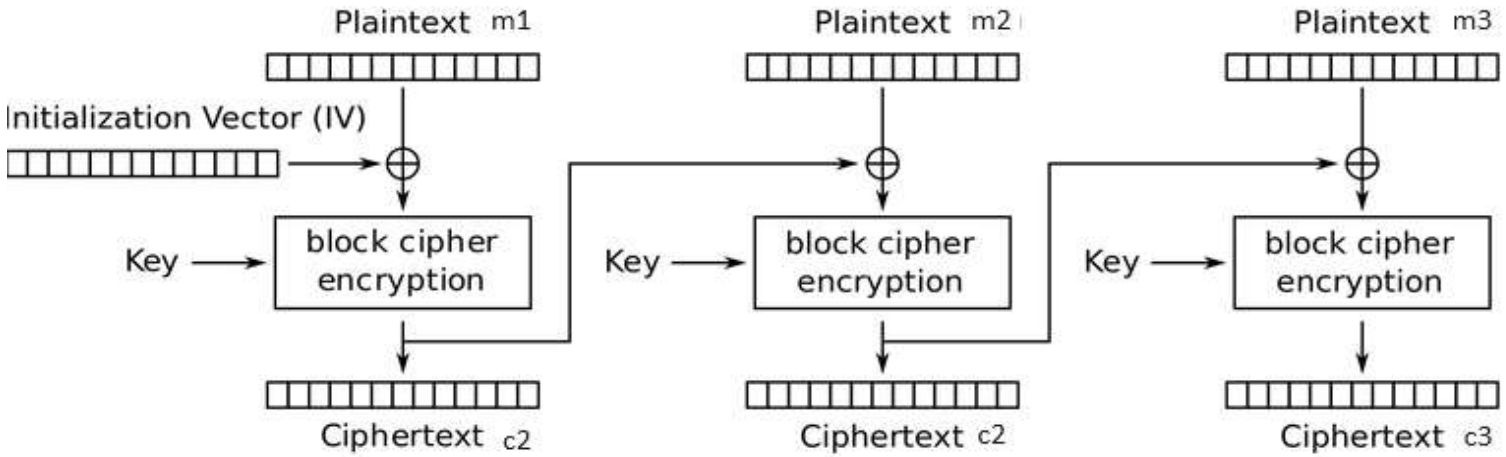


Diagram shows how block cipher works in CBC mode
(CBC = cipher block chaining)



Principle:

Message m is divided into 128 bit blocks m1, m2, m3, …

Key k has 128 bits in AES128. Cipher consists of blocks c1, c2, c3, ...

In CBC mode each output block is taken as 3rd input in calculation of next cipher block.

$$c_n = AES(m_n, k, c_{n-1})$$

Requirements for block ciphers

- Block cipher should have both software and hardware (chip) implementations (AES has both)
- Algorithm should be realiable and fast
- It can be applied to both data transfer and files
- Algorithm should be public and transparent (Kerckhoff principle)

- Key lenght should be at minimum 128 bits

Shannon's diffusion and confusion

The operation of block ciphers is based on two principles of Claude Shannon: diffusion and confusion

Diffusion: (dependence of cipher and message)
Changing one bit of message block should change in average half of the output bits. (=> the output bits should equally depend on all input bits)

Confusion: (dependence of cipher and key)
Every bit of cipher should depend on all parts of the key. The change of one bit of the key, should change the output completely (not just part of it)

The ciphertext should be random and fulfill the properties of pseudorandom bit sequence (explained later)

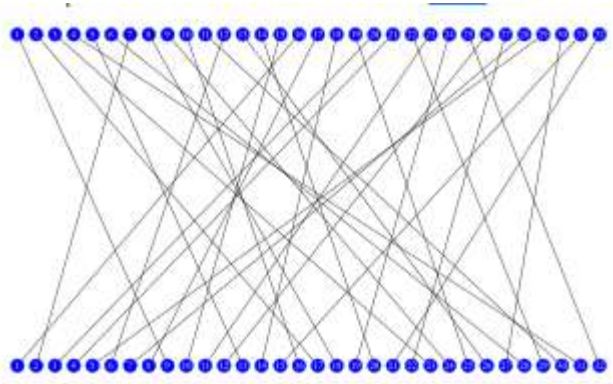
Permutation – Substitution networks

In practice the principles of diffusion and confusion are mostly implemented usign a network of permutations and substitutions.

Substitution = replacing certain bit patterns with image patterns, traditionally using substitution tables called SBOX:es. Below a table of Sbox nr 6 of DES block cipher. In example bit sequence 011011 is mapped to 1001

S_5		Middle 4 bits of input															
		0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
Outer bits	00	0010	1100	0100	0001	0111	1010	1011	0110	1000	0101	0011	1111	1101	0000	1110	1001
	01	1110	1011	0010	1100	0100	0111	1101	0001	0101	0000	1111	1010	0011	1001	1000	0110
	10	0100	0010	0001	1011	1010	1101	0111	1000	1111	1001	1100	0101	0110	0011	0000	1110
	11	1011	1000	1100	0111	0001	1110	0010	1101	0110	1111	0000	1001	1010	0100	0101	0011

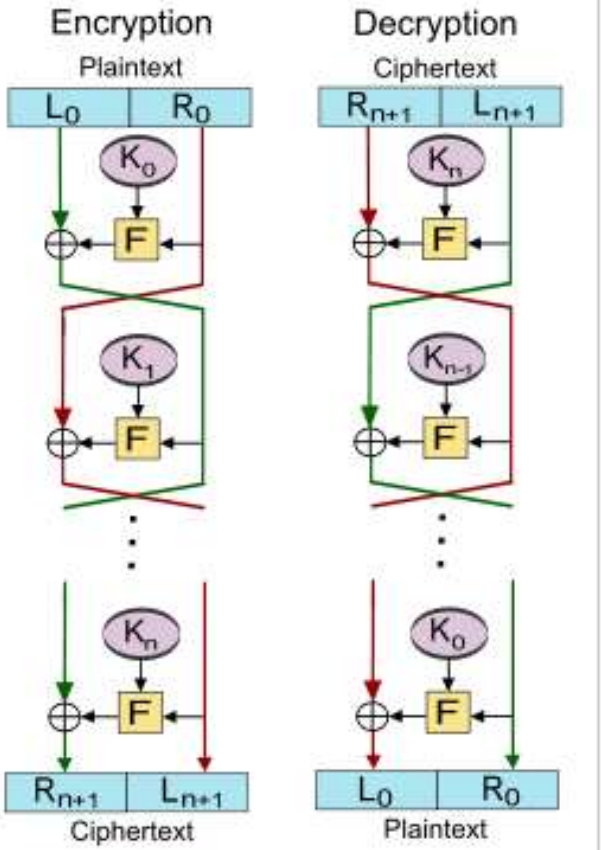
Permutation = reordering bits of a block using permutation table. Figure shows how 64 bits are rearranged in P-box of DES block cipher.



Rounds: Permutation- Substitution loop is run several times,

(f.e in DES algorithm 16 times) to remove all regularities in the cipher bit stream.

Feistel network



Many block ciphers, such as DES and Blowfish utilize structures known as *Feistel ciphers*

IBM engineer Feistel created PS network in 1960's using permutations and SBOX:s (substitution tables).
Feistel network appeared originally in DES block cipher, it was later reused in Blowfish cipher and later in **Kasumi , which was recently a standard block cipher in 3G and 4G** (Today Kasumi is replaced by AES and SNOW3G)

Feistel algorithm has 16 rounds. Each round uses different subkey Ki derived from the master key K

Each round starts with initial permutation and ends with final permutation (called P-BOX:s)

Between permutations there are eight S-BOX:s, which obscure the connection between the key and the ciphertext (Shannon's confusion)

AES encryption standard

In late 1990's NIST organized a competition to choose a new block cipher standard. There were 6 finalist, from which the Rijndael algorithm was chosen as the new standard in 2001.

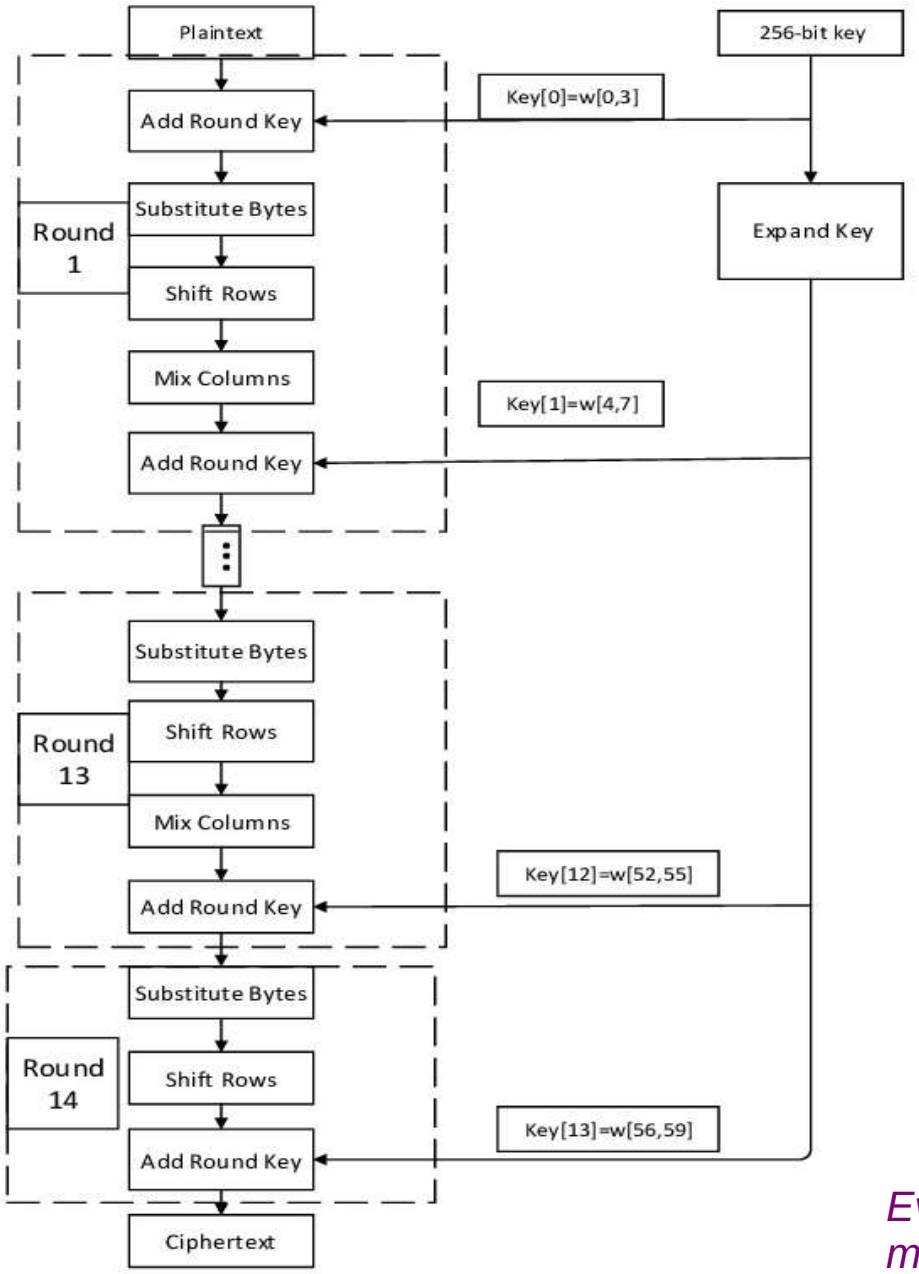
The winner got a new name *Advanced Encryption Standard* AES.

Advantages of AES

* It has lots of versions AES128, AES196, AES256, AES512. With flexible key lengths it answer the challenges of increasing computing power

* The security is based on provable, transparent mathematics

It is generally thought that AES will be used as the standard block cipher for a long time in the forseeable future.



AES256 uses 256 bit key

Algorithm has 14 rounds all of which use different subkeys derived from the primary key

Each round has many phases, which implement diffusion and confusion of bits

Phases of rounds are:
Substitute bytes (Sbox)
Shift rows
Mix columns
Add round key

The finest part, Sboxes use Theory of Galois Fields. The algorithm in Sboxes was invented in 1991 by Finnish cryptographer Kaisa Nyberg

Evariste Galois (1811-1832) was a great French mathematician who died at the age of 21

Steps of AES rounds in pictures

For mathemically oriented students the multiplication rule of Galois Fields is added below.

Multiplication a*b in Galois Field 2^8.
Example: a=11010100, b = 01101001

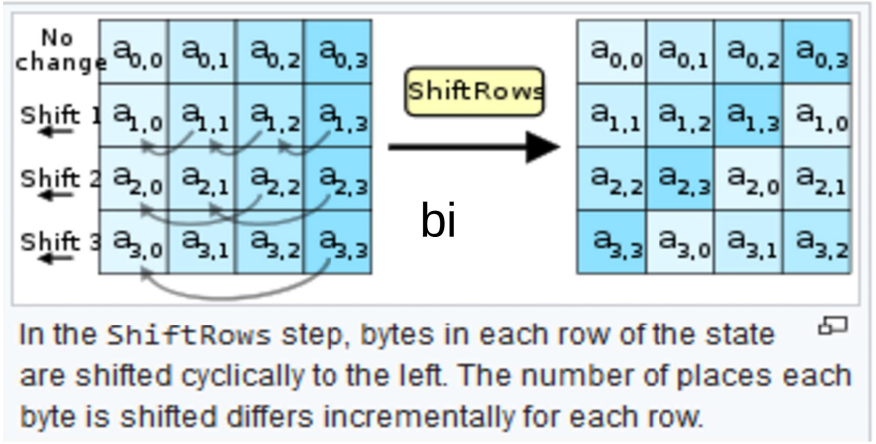
1) a and b are intepreted as polynomials
$a = x^7+x^6+x^4+x^2$, $b=x^6+x^5+x^3+1$'

2) $a*b = x^{13} + 2x^{12} + x^{11} + 2x^{10} + 2x^9 + x^8 + 3x^7 + x^6 + x^5 + x^4 + x^2$
Converting coefficients mod 2
$a*b = x^{13} + x^{11} + x^8 + x^7 + x^6 + x^5 + x^4 + x^2$

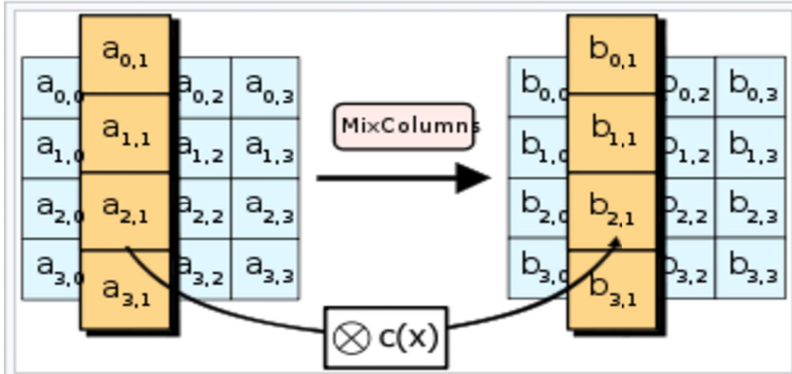
3) Finally divide stage 2 result with $x^8 + x^4 + x^3 + x + 1$. The final result of a*b is the remainder polynomial $x^6 + x^5 + x^4 + x^3 + x$ corresponding binary number 01111010

Thus 11010100 x 01101001=01111010

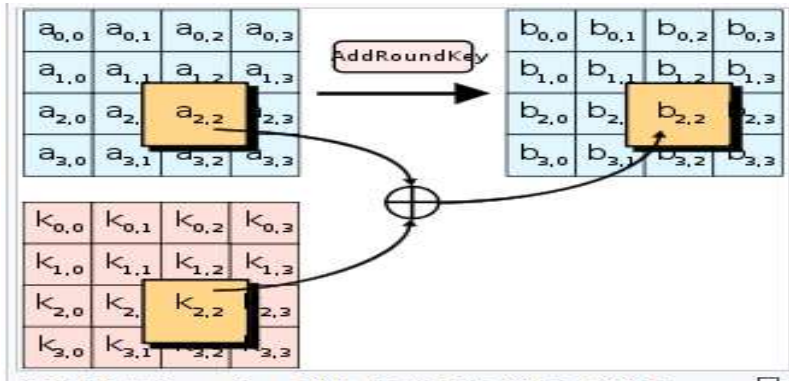
AES uses multiplication in GF(2^8) as in the example. The divisor polynomial $x^8 + x^4 + x^3 + x + 1$ used as "modulus" is called "Field Irreducible polynomial". Idea is analogical with multiplication of integers mod p, where p is prime modulus



In the ShiftRows step, bytes in each row of the state are shifted cyclically to the left. The number of places each byte is shifted differs incrementally for each row.



In the MixColumns step, each column of the state is multiplied with a fixed polynomial $c(x)$.



In the AddRoundKey step, each byte of the state is combined with a byte of the round subkey using the XOR operation (⊕).

Modes of operation

1) ECB Electronic Codebook

Message blocks are encrypted independently => Identical message blocks have identical cipher blocks. This is why ECB mode is not secure and not used in block ciphers

2) CBC Cipher Block Chaining (used in TLS until 2017)

Each cipher block functions as extra input of encryption of next block. The ciphers of first block affects all remaining blocks. This increases security.

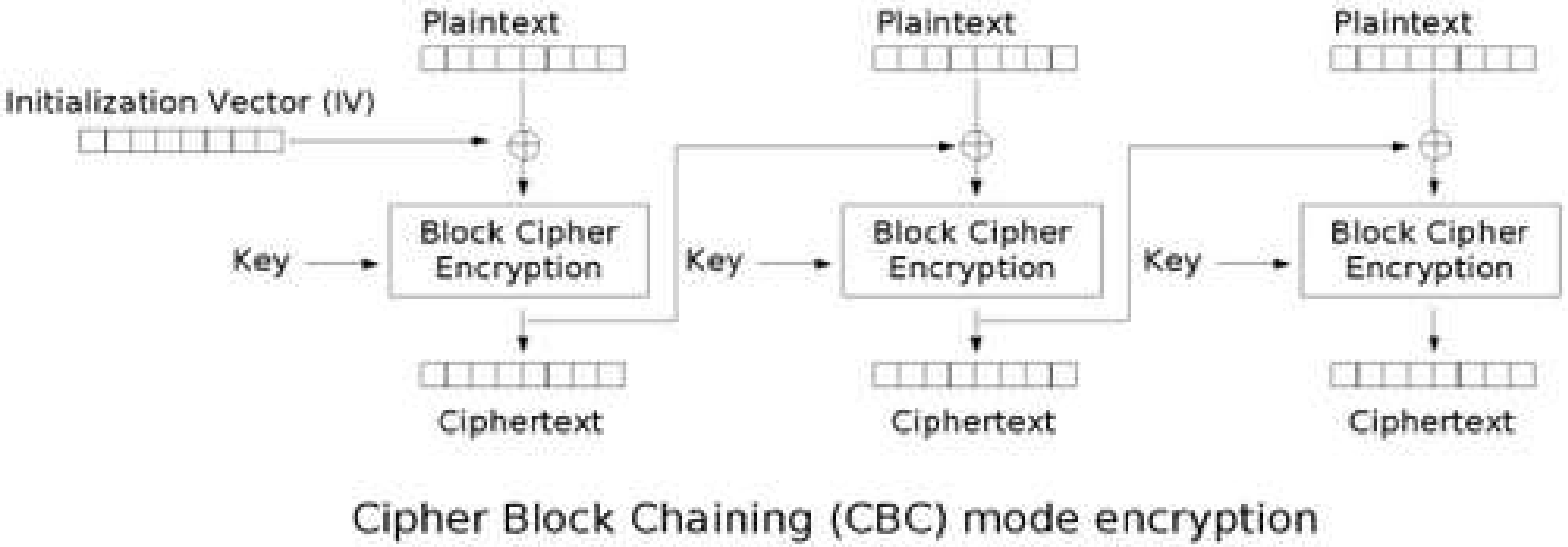
The following line from 2017 describing algorithms of Nordea Bank's TLS connections

3) **GCM Galois Counter Mode** (today standard in TLS)

In 2023 Nordea Bank's TLS uses AES in GCM mode

About AES_CBC and AES_GCM modes

One can still find f.e email systems using AES_CBC .



Cipher Block Chaining (CBC) mode encryption

Data is encrypted in 128 bit blocks. Input of each encryption are message, key, and the cipherblock of previous encryption (or IV). Every cipherblocks affects all next cipher blocks, which makes the encryption stronger against attacks and makes manipulations of some part of the cipher during transmission impossible..

GCM – mode (”Galois Counter Mode”)

has replaced CBC mode in net banks providing stronger protection of integrity and authentication of sender.

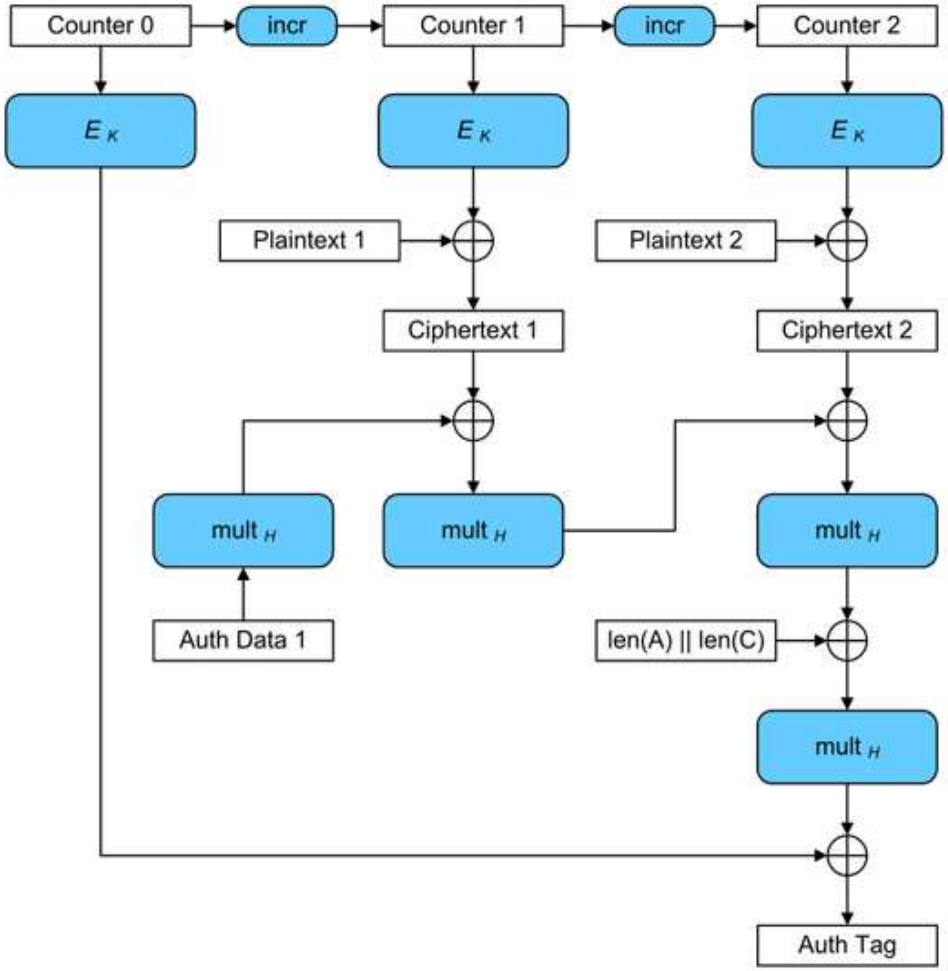
Tekniset tiedot

Yhteys salattu (TLS_ECDHE_ECDSA_WITH_AES_256_GCM_SHA384, 256-bittinen avain, TLS 1.2)

In GCM- mode counter value (1,2,…) is encrypted with AES with key K. The output (128 bits) is XOR:ed with the first message block m1 to produce the first cipher block c1.

Authentication data (senders personal data + previous cipher blocks) is collected to a 384 hash value using Galois Field multiplication rule.

The goal is to make it impossible to attack the encryption and secure the integrity and authentication.



Finnish governments instructions for block ciphers

National security classification	IV	III	II	I
Block cipher	AES 128 Serpent(128)	AES 196 Serpent(196)	AES 256 Serpent(256)	none *)

*) Top secret documents (class 1) should be kept only in paper form.

In practice the government uses AES encryption

Source: Cyper security center of Finland

Is AES post-quantum secure?

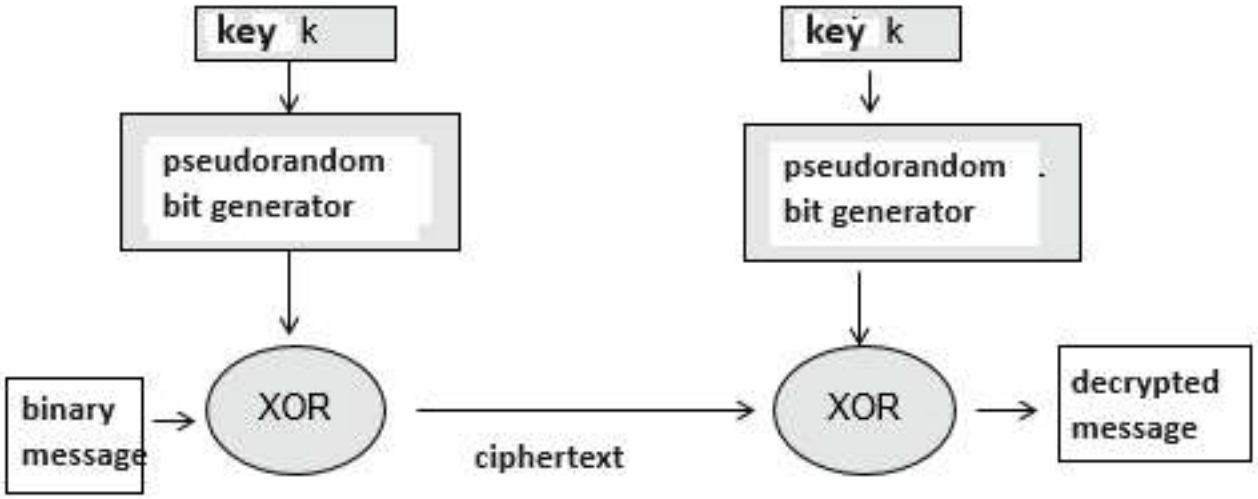
Quantum computers will be ”game changers” in cryptography.

According to today's knowledge AES128 will no longer be secure in post-quantum world, but AES256 and AES512 still provide adequate security. *)

*) *There is a quantum algorithm called Grover’s algorithm, which could crack AES128 key in a few seconds using 128 cubit quantum computers.*

However Grover’s algorithm applied to AES256 is not able to crack the 256 bit key. The security of AES256 remains at the level that corresponds AES128 in the present pre-quantum world.

Syncronous stream ciphers



The **encryption algorithm A5** (1991) of GSM (G2) phones was a syncronous stream cipher. It imitates the idea of One Time Pad.

GSM phones have a micro circuit which produces a pseudorandom bit sequence, which is added with XOR to the binary message.

GSM call would be perfectly safe, if the key sequence would be truly random, which it is not: key sequence looks random, but it is deterministic.

GSM calls are not secure, encyption of GSM calls can be broken. GSM network has been replaced by 3G, 4G and 5G networks, which are more secure.

Definition of pseudorandomness

PN- sequence (PN = ”pseudo noise”) or pseudorandom bit sequence is a key -based, deterministic bit sequence, which satisfies three properties defined by **Samuel W Golomb** in 1950’s..

PN sequence properties:

G1. Sequence must have 50% ones 50% zeros

G2. Relative frequencies of blocks repeating same bit (0 or 1)

0**1**0	**or**	1**0**1	single bit block
0**11**0		1**00**1	two identical bits block
0**111**0		1**000**1	three identical bits block
0**1111**0		1**0000**1	four identical bits block

have ratios **1/2 , 1/4 , 1/8 , 1/16**

G3. When a sequence is rotated and the rotated sequence is compared with the original, the number of incidences and disincidences should be approximately equal. This property is measured by so called autocorrelation coefficient, which should be close to zero for all rotations of sequence, which means that the sequence does not have internal periodicity.

”Cryptographically secure random number generation”

Secure communication needs **”cryptographically secure” pseudorandom number generation (CSPRNG)** for example in creating symmetric keys for AES.
(All pseudorandom numbers are not cryptographically secure).

Some CSPRNG methods

1. **Enthropy source method**. Unpredictable data stream is collected mainly from computers operating system (key inputs, time stamps, interruptions, sensors, ….). This data is hashed to create random numbers.

2. ”**Streching entropy”.** Data obtained from 1st method is limited, if lots of random numbers are needed at once. One can extend the quantity of random numbers using for example AES in counter mode. Each output is a random number..

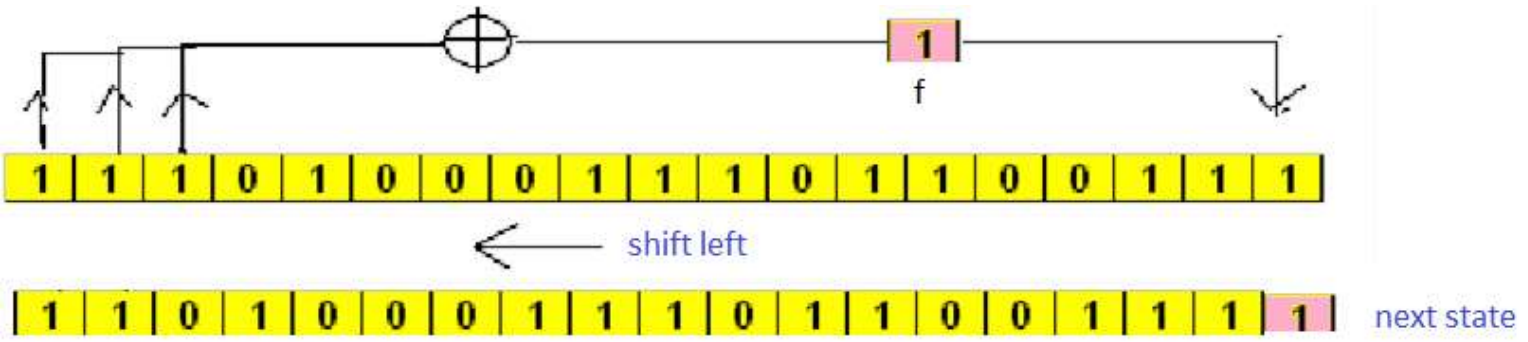
Random key generation is critical issue for safe communication. A non-safe randon number generator has been deliberately used for eavedropping communication.
”The Guardian and The New York Times have reported in 2013 that the National Security Agency (NSA) inserted a backdoor into a pseudorandom number generator (PRNG) of NIST SP 800-90A which allows the NSA to readily decrypt material that was encrypted with the aid of Dual EC DRBG. Both papers report[30][31] that, as independent security experts long suspected,[32] the NSA has been introducing weaknesses into CSPRNG standard 800-90; this being confirmed for the first time by one of the top secret documents leaked to the Guardian by Edward Snowden.” (Wikipedia)

LFSR - linear feedback shift register

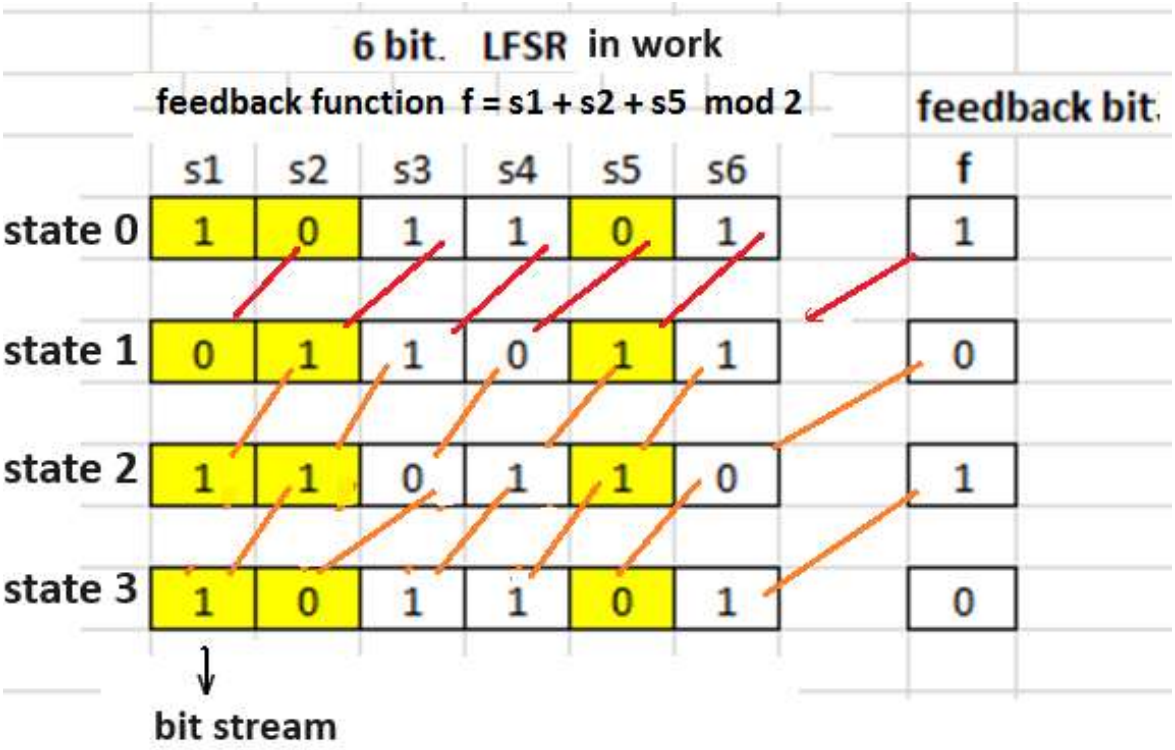
= microcircuit traditionally used in RAND of pocket calculators and later in GSM phones

* LFSR is n bit register which can be used to produce random numbers.
During every clock pulse the bits shift one step to the left.
The new rightmost bit, called the feedback bit, is calculated using XOR addition from specific bit values just before the shift.

Feedback bit f = 1+ 1+1 mod 2 = 1



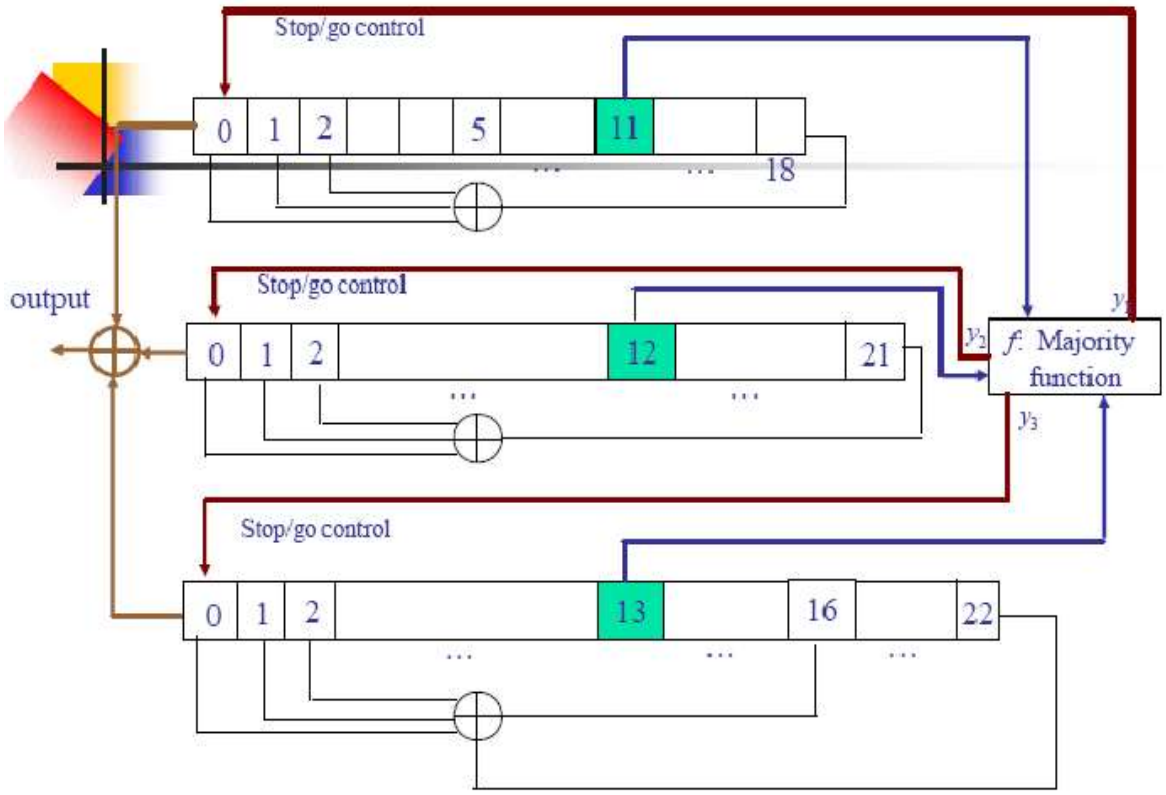
Example of LFSR in operation



Feedback bit is calculated using formula
$f = s_1+s_2+s_5 \pmod 2$
before bits are shifted to left.

LFSR registers leftmost S1 – bits form a sequence, which fulfills the three properties of definition of pseudorandomness.

GSM phones A5 PRNG -generator



Majority function: I all marked green bits have identical values, all registers move forwards. In other cases those two move, which have identical values .

GSM – phone has 3 LFSR-registers of lenghts 19, 22 and 23 bits

Initial state (64 bits = 19+22+23) is the symmetric key

At each clock pulse circuit generates one pseudo random bit (a XOR-sum of zero bits)

To increase security a majority function f is added to the circuit. It is calculated from bits marked green in the picture, which of the three registers are shifted during one clock pulse and which remain as they are.

A5 bit stream fulfills pseudorandomness postulates G1, G2, G3

```
0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1,
1, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 1,
1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0,
1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1,
1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0,
1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1,
1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0,
0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 0, 1,
1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0,
1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1,
0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1,
1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 0,
1, 1, 0, 1, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0,
1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1,
1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0,
```

GSM encryption: The key stream and digitalized voice are added using XOR to produce the encrypted signal

Operators server has a similar PRNG and same symmetric initial state Operator decrypts signal.using XOR. Signal is encrypted only in air, not in the cable.

GSM security

1. Key in A5 encryption is 64 bits, which is not secure.
2. Key sequence is periodic, effective period is $4/3*2^{23}$
3. GSM signal is encrypted only between the mast and the phone (not in cables)
4. GSM is not secure against professional hackers

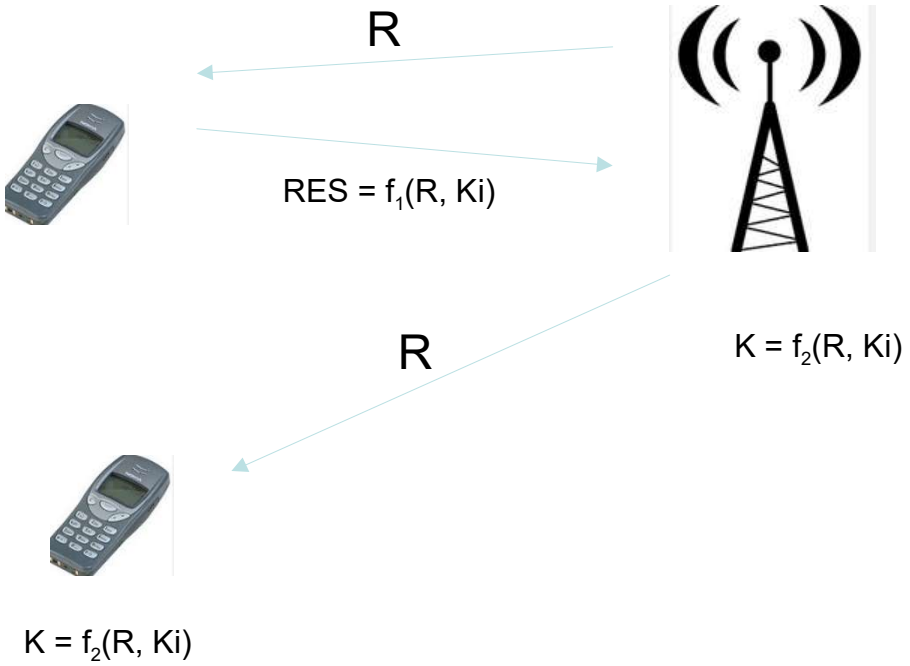
Authentication and key agreement in GSM

Authentication A3. Operator sends to the phone a random challenge number R.
Phone calculates and sends a response RES from R and SIM-key.

Operator calculates RES in the same way. If there is a match, the phone is authenticated.

Key agreement A8. Operator sends the phone a random R. The phone calculates a 64 bit key K from R and SIM-key.

The key K is the initial state for the three LFSR -registers



3G, 4G, 5G network encryption

3G and 4G use block ciphers: AES128 of SNOW3G

- key length = 128 bits
- two way authentication is used: both phone and the mast must authenticate themselves in the beginning of communication

5G encryption uses 256 bit versions of AES and SNOW3G