

## Lecture 3: Introduction to cryptography. Secret key cryptography. Hash functions

Course instructors: Alexey Frolov and Yury Yanovich

Teaching Assistant: Stanislav Kruglik

Technical Assistants: Anton Glebov and Evgeny Marshakov

November 6, 2018

- 1 Introduction to cryptography
- 2 Information-theoretic security
  - Measure of information
  - Perfect secrecy
- 3 Secret key cryptography
- 4 Hash functions

Alice wants to send a private message to Bob over a public network

- ▶ What if someone intercepts and reads this message?  
**(Confidentiality)**
- ▶ What if someone intercepts and alters this message?  
**(Integrity)**
- ▶ What if someone pretending to be Alice forges a message and sends it to Bob? **(Authentication)**
- ▶ What if Alice denies sending the message? **(Non-repudiation of origin, Digital Signature)**
- ▶ What if Bob denies receiving the message?  
**(Non-repudiation of destination)**

- ▶ Encryption scheme (for privacy)
  - ▶ functions to encrypt, decrypt data
  - ▶ key generation algorithm
- ▶ Symmetric key (private key) vs. asymmetric key (public key)
  - ▶ **Public key**: publishing public key **key** does not reveal secret key  $\text{key}^{-1}$
  - ▶ **Private key**: more efficient, generally  $\text{key} = \text{key}^{-1}$
- ▶ Hash function, MAC (for integrity, authenticity)
  - ▶ A **hash function** maps any input to a short hash; ideally, **no collisions**
  - ▶ MAC (keyed hash) used for message integrity
- ▶ Signature scheme (for integrity, authenticity)
  - ▶ Functions to sign data, verify signature

- ▶ Secret Key (also known as private, single, symmetric key) existing for more than 1000 years
- ▶ Public Key (also known as two key, asymmetric key) since 1974, both secret key and public key systems are in use and competing with each other

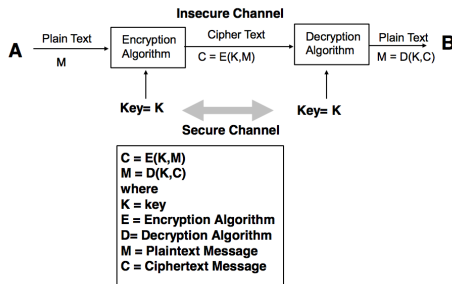
- ▶ Used to make content unreadable by all but the intended receivers

$E(\text{plaintext}, \text{key}) = \text{ciphertext}$

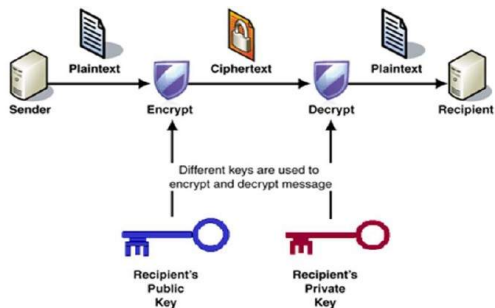
$D(\text{ciphertext}, \text{key}) = \text{plaintext}$

- ▶ Algorithm is public, key is private
- ▶ Block (input fixed blocks) vs. Stream Ciphers (stream of input)

# Secret key cryptosystem



# Public key cryptosystem





- ▶ Information-theoretic security (unconditional)
  - ▶ A ciphertext  $C$  (encryption) leaks absolutely no information about the message  $M$  (plaintext).
  - ▶ Even attackers with infinite computational resources cannot learn any information of  $M$  from the ciphertext.
  - ▶ Requires same ciphertext distribution for all plaintexts, i.e.  
$$\Pr[C = c | M = m] = \Pr[C = c]$$
  - ▶ Example: one-time pads
  - ▶ Also a quantitative measure in statistical databases.
- ▶ Computational security
  - ▶ A ciphertext may leak information about the message  $M$ , but a lot of computation is needed to extract this information.
  - ▶ Adopted by all practical schemes.

- 1 Introduction to cryptography
- 2 Information-theoretic security
  - Measure of information
  - Perfect secrecy
- 3 Secret key cryptography
- 4 Hash functions

How to measure the randomness of R.V.  $X$ ?

Shannon entropy  $H(p_1, p_2, \dots, p_n)$  is characterized by a small number of criteria ( $p_i = P(x_i)$ ).

- ▶ **Continuity.** Changing the values of the probabilities by a very small amount should only change the entropy by a small amount.
- ▶ **Symmetry.** E.g.  $H(p_1, p_2, \dots, p_n) = H(p_2, p_1, \dots, p_n)$ .
- ▶ **Maximum.** The measure should be maximal if all the outcomes are equally likely (uncertainty is highest when all possible events are equiprobable). For equiprobable events the entropy should increase with the number of outcomes.
- ▶ **Additivity.** The amount of entropy should be independent of how the process is regarded as being divided into parts.

$$H\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = H\left(\frac{b_1}{n}, \dots, \frac{b_k}{n}\right) + \sum_{i=1}^k \frac{b_i}{n} H\left(\frac{1}{b_1}, \dots, \frac{1}{b_k}\right)$$

Any definition of entropy satisfying these assumptions has the form

$$H(X) = -c \sum_{x \in \mathcal{X}} P(x) \log P(x) = c \mathbb{E}[\log(1/P(X))],$$

where  $c$  is a constant corresponding to a choice of measurement units. We agree that  $0 \log 0 = 0$ .

$\log_2 \leftrightarrow$  bits

$\ln \leftrightarrow$  nats

$\log_{256} \leftrightarrow$  bytes

$\log \leftrightarrow$  arbitrary units, base always matches exp

$$H(X) \geq 0$$

## Definition

The *joint* entropy  $H(X, Y)$  can be defined as

$$H(X, Y) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P(x, y) \log P(x, y) = -\mathbb{E}[\log P(x, y)].$$

## Definition

The *conditional* entropy  $H(Y|X)$  can be defined as

$$\begin{aligned} H(X, Y) &= \sum_{x \in \mathcal{X}} P(x) H(Y|x) \\ &= - \sum_{x \in \mathcal{X}} P(x) \sum_{y \in \mathcal{Y}} P(y|x) \log P(y|x) \\ &= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P(x, y) \log P(y|x) \\ &= \mathbb{E}[\log P(Y|X)]. \end{aligned}$$

Entropy is a measure of uncertainty of R.V. Relative entropy is a measure of distance in between two distributions  $P(x)$  and  $Q(x)$ .

## Definition

The relative entropy or Kullbak–Leibler distance in between p.m.f.'s  $P(x)$  and  $Q(x)$  is defined as

$$D(P||Q) = \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)} = \mathbb{E} \left[ \frac{P(X)}{Q(X)} \right].$$

We use convention  $0 \log \frac{0}{q} = 0$  and  $p \log \frac{p}{0} = \infty$ .

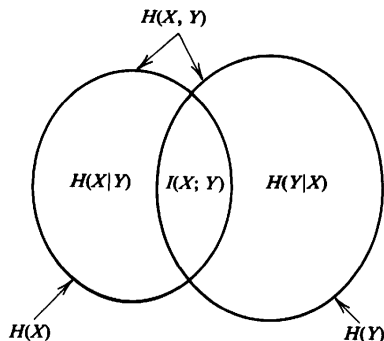
Is it a metric?



## Definition

$$\begin{aligned} I(X; Y) &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P(x, y) \log \frac{P(X, Y)}{P(X)P(Y)} \\ &= D(P_{X,Y} \| P_X P_Y) \\ &= \mathbb{E} \left[ \log \frac{P(X, Y)}{P(X)P(Y)} \right] \end{aligned}$$

$$I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X).$$



- ▶ Consider a secret key cryptosystem with the same alphabet for ciphertext and plaintext;
- ▶ Plaintext  $M$  and key  $K$  are random variables, i.e. we are given a joint distribution  $P(m, k)$

Properties:

- ▶  $H(M|C, K) = 0$ ;
- ▶  $H(C|M, K) = 0$ ;
- ▶ Assumption:  $I(M, K) = 0$ ;
- ▶  $P(c) = \sum_{(k,m:E(m,k)=c)} P(k)P(m)$ ;
- ▶  $P(c|m) = \sum_{(k:D(c,k)=m)} P(k)$ ;
- ▶  $P(m|c) = \frac{P(c|m)P(m)}{P(c)}$ ;

Main condition:

$$I(M; C) = 0.$$

Corollary:

$$H(M|C) = H(M).$$

$$H(M) = H(M|C) \leq H(M, K|C) = H(K|C) + H(M|C, K).$$

As  $H(M|C, K) = 0$  and  $H(K|C) \leq H(K)$  we obtain

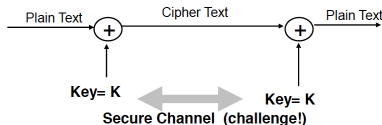
$$H(M) \leq H(K).$$

Let us denote the average length of a plaintext by  $L(M)$  and the average length of a key by  $L(K)$ , we have

$$L(M) \leq L(K).$$

# Vernam One-time Pad

- ▶ Ultimate cipher
- ▶ impractical for most solutions
- ▶ requires a perfectly random key as large as the message
- ▶ the key cannot be reused
- ▶ known plaintext reveals the portion of the key that has been used on it, but does not reveal anything about the other bits of the key



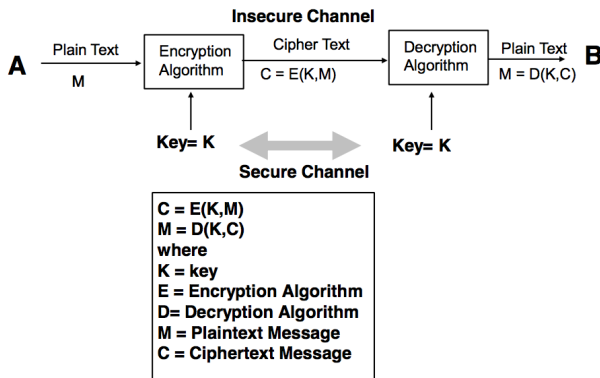
- ▶ Share a random key  $k$
- ▶ Encrypt plaintext by XOR ( $\oplus$ ) with sequence of bits  
 $E(k, m) = k \oplus m$  (bit-by-bit)
- ▶ Decrypt ciphertext by XOR ( $\oplus$ ) with same bits  
 $D(k, c) = k \oplus c$  (bit-by-bit)
- ▶ Advantages
  - ▶ Very efficient
  - ▶ This is an information-theoretically secure cipher
- ▶ Disadvantages
  - ▶ Key is as long as the plaintext
  - ▶ How does sender pass key to receiver securely?

Idea for stream cipher: use pseudo-random generators for key



- 1 Introduction to cryptography
- 2 Information-theoretic security
  - Measure of information
  - Perfect secrecy
- 3 Secret key cryptography
- 4 Hash functions

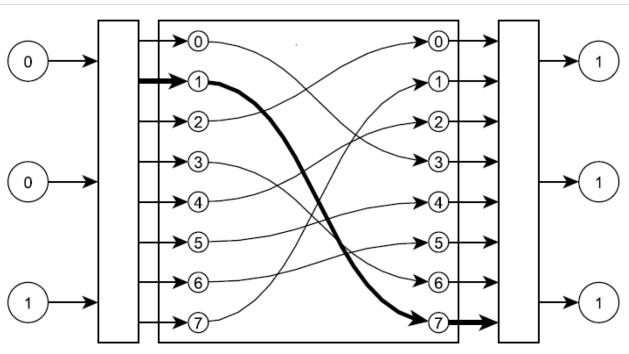
# Secret key cryptosystem



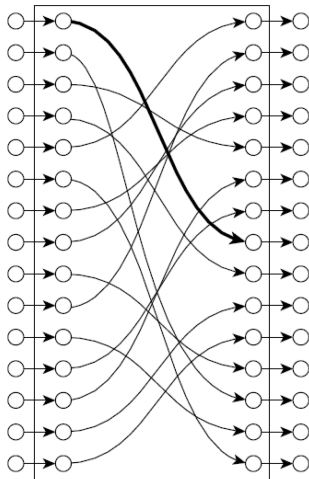
- ▶ operates on fixed-length groups of bits, called a block
- ▶ unvarying transformation that is specified by a symmetric key

Important elementary components in the design of many cryptographic protocols!

- ▶ The modern design of block ciphers is based on the concept of an iterated product cipher.
- ▶ In his seminal 1949 publication, Communication Theory of Secrecy Systems, Claude Shannon analyzed product ciphers and suggested them as a means of effectively improving security by combining simple operations such as substitutions and permutations.
- ▶ Iterated product ciphers carry out encryption in multiple rounds, each of which uses a different subkey derived from the original key.
- ▶ One widespread implementation of such ciphers, named a Feistel network after Horst Feistel, is notably implemented in the DES cipher. Many other realizations of block ciphers, such as the AES, are classified as substitutionpermutation networks.



- ▶ Big substitution box (a set of S-boxes, one S-box per each key)
- ▶ Drawback: memory consumption

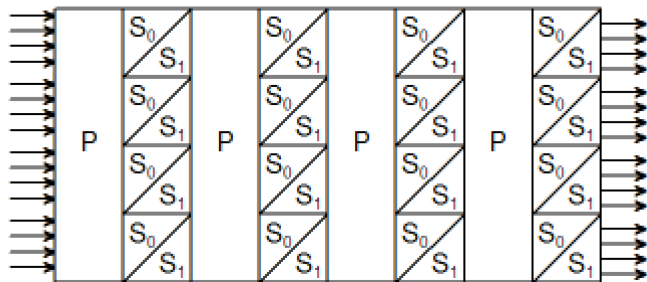


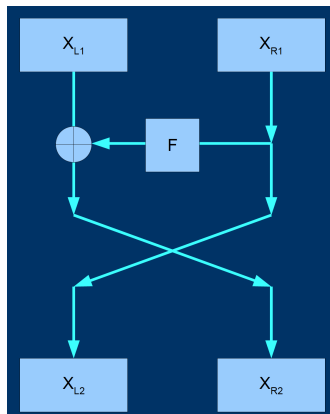
- ▶ Simple implementation
- ▶ Drawback: not secure



Let's combine S- and P-boxes. When the number of “layers” is sufficient the resulting SP-network works like a big S-box.

# Lucifer cipher (Feistel, 1973)

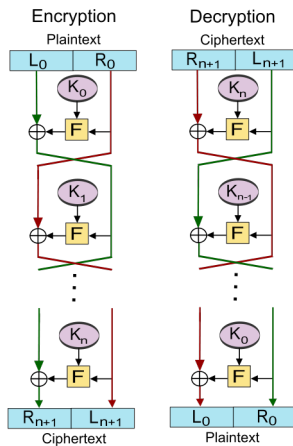




$$L_{i+1} = R_i$$

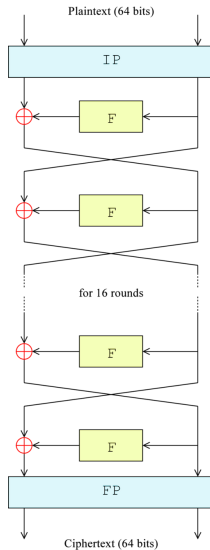
$$R_{i+1} = L_i \oplus F(R_i, K_i).$$

# Feistel cipher



- ▶ First published: 1975
- ▶ Standardized: 1977 (Federal Information Processing Standard, FIPS)
- ▶ Block size: 64 bits
- ▶ Key size: 56 bits (+8 parity bits)
- ▶ Structure: Balanced Feistel network
- ▶ Now considered to be insecure for many applications (due to small key size)

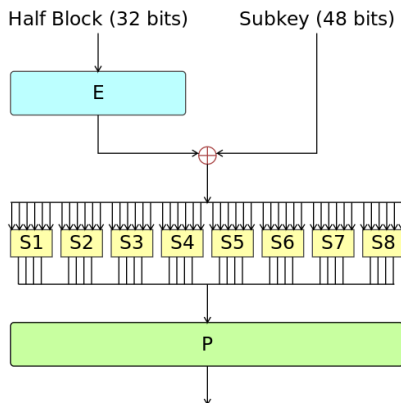
# Overall structure



- ▶ 16 identical stages of processing, termed rounds.
- ▶ initial and final permutation, termed IP and FP, which are inverses (IP "undoes" the action of FP, and vice versa).
- ▶ IP and FP have no cryptographic significance

# Feistel function (F)

The F-function,



operates on half a block (32 bits) at a time and consists of four stages.



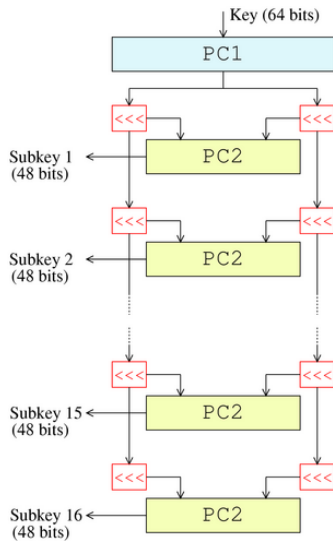
Stage 1 (Expansion): the 32-bit half-block is expanded to 48 bits using the expansion permutation, denoted  $E$  in the diagram, by duplicating half of the bits. The output consists of eight 6-bit ( $8 \times 6 = 48$  bits) pieces, each containing a copy of 4 corresponding input bits, plus a copy of the immediately adjacent bit from each of the input pieces to either side.

Stage 2 (Key mixing): the result is combined with a subkey using an XOR operation. Sixteen 48-bit subkeys (one for each round) are derived from the main key using the key schedule.

Stage 3 (Substitution): after mixing in the subkey, the block is divided into eight 6-bit pieces before processing by the S-boxes, or substitution boxes. Each of the eight S-boxes replaces its six input bits with four output bits according to a non-linear transformation, provided in the form of a lookup table. The S-boxes provide the core of the security of DES – without them, the cipher would be linear, and trivially breakable.

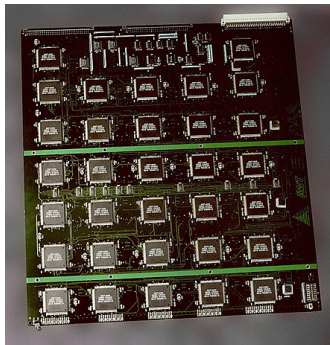
Stage 4 (Permutation): finally, the 32 outputs from the S-boxes are rearranged according to a fixed permutation, the P-box. This is designed so that, after permutation, the bits from the output of each S-box in this round are spread across four different S-boxes in the next round.

# Key schedule



- ▶ Initially, 56 bits of the key are selected from the initial 64 by Permuted Choice 1 (PC-1)
- ▶ The 56 bits are then divided into two 28-bit halves; each half is thereafter treated separately.
- ▶ In successive rounds, both halves are rotated left by one or two bits (specified for each round), and then 48 subkey bits are selected by Permuted Choice 2 (PC-2) – 24 bits from the left half, and 24 from the right.
- ▶ The key schedule for decryption is similar – the subkeys are in reverse order compared to encryption. Apart from that change, the process is the same as for encryption. The same 28 bits are passed to all rotation boxes.

# Brute-force attack



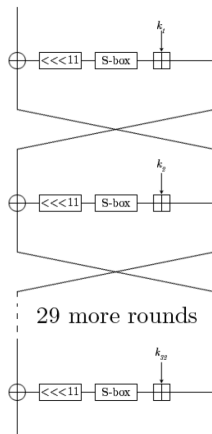
US \$250,000 DES cracking machine contained 1,856 custom chips and could brute-force a DES key in a matter of days.

DES exhibits the complementation property, namely that

$$E_K(P) = C \Leftrightarrow E_{\bar{K}}(\bar{P}) = \bar{C}.$$

where  $\bar{x}$  is the bitwise complement of  $x$ . The complementation property means that the work for a brute-force attack could be reduced by a factor of 2 (or a single bit) under a chosen-plaintext assumption.





- ▶ 1989 year
- ▶ Key size: 256 bits
- ▶ Block size: 64 bits
- ▶ Number of rounds: 16 / 32
- ▶ 8 S-boxed (4-bit)

## DES:

- ▶ 1977 year
- ▶ 56 bit key
- ▶ 64 bit block
- ▶ 16 rounds
- ▶ replaced with 3DES in 1990-th
- ▶ replaced with AES in 2001

## GOST:

- ▶ 1989 year
- ▶ 256 bit key
- ▶ 64 bit block
- ▶ 16/32 rounds
- ▶ currently used

Blowfish, Camellia, CAST, CIPHERUNICORN, CLEFIA, DEAL, DFC, FEAL, IDEA, KASUMI, Knufu, LOKI, MacGuffin, MAGENTA, MARS, MISTY, Raiden, RC2 / RC5 / RC6, RTEA, SEED

- ▶ Avalanche effect
- ▶ S-boxes
- ▶ Efficient implementation

- ▶ 1998 year
- ▶ Block size: 64 bits
- ▶ Key sizes: 56/112/168 bits
- ▶ 48 rounds

Main idea: increasing the key size of DES to protect against brute-force attacks, without the need to design a completely new block cipher algorithm.

Naive approach:

$$E_{K_2}(E_{K_1}(\text{plaintext}))$$

given a known plaintext pair  $(x, y)$ , such that  $y = E_{K_2}(E_{K_1}(x))$ , one can recover the key pair  $(K_1, K_2)$  in  $\sim 2^n$  steps



The encryption algorithm is:

$$\text{ciphertext} = E_{K_3}(D_{K_2}(E_{K_1}(\text{plaintext}))).$$

Decryption is the reverse:

$$\text{plaintext} = D_{K_1}(E_{K_2}(D_{K_3}(\text{ciphertext})))$$

# What about keys?

- **Keying option 1.** All three keys are independent. Sometimes known as 3TDEA or triple-length keys. This is the strongest, with  $3 \times 56 = 168$  independent key bits.

# What about keys?

- ▶ **Keying option 1.** All three keys are independent. Sometimes known as 3TDEA or triple-length keys. This is the strongest, with  $3 \times 56 = 168$  independent key bits.
- ▶ **Keying option 2.**  $K_1$  and  $K_2$  are independent, and  $K_3 = K_1$ . Sometimes known as 2TDEA or double-length keys. This provides a shorter key length of 112 bits and a reasonable compromise between DES and Keying option 1, with the same caveat as above.

# What about keys?

- ▶ **Keying option 1.** All three keys are independent. Sometimes known as 3TDEA or triple-length keys. This is the strongest, with  $3 \times 56 = 168$  independent key bits.
- ▶ **Keying option 2.**  $K_1$  and  $K_2$  are independent, and  $K_3 = K_1$ . Sometimes known as 2TDEA or double-length keys. This provides a shorter key length of 112 bits and a reasonable compromise between DES and Keying option 1, with the same caveat as above.
- ▶ **Keying option 3.** All three keys are identical, i.e.  $K_1 = K_2 = K_3$ . This is backward compatible with DES, since two operations cancel out. ISO/IEC 18033-3 never allowed this option, and NIST no longer allows it.

- ▶ Competition (1998), organized by NIST
- ▶ Requirements: block cipher, 128 bit block, key sizes – 128, 192 and 256 bits.

Participants: CAST-256, CRYPTON, DEAL, DFC, E2, FROG, HPC, LOKI97, MAGENTA, MARS, RC6, Rijndael, SAFER+, Serpent, Twofish

First round: CAST-256, DFC, E2, LOKI-97, MAGENTA, MARS, Rijndael, SAFER+, Serpent



- ▶ Unlike its predecessor DES, AES does not use a Feistel network.
- ▶ AES operates on a  $4 \times 4$  column-major order matrix of bytes, termed the *state*. Most AES calculations are done in a particular finite field.

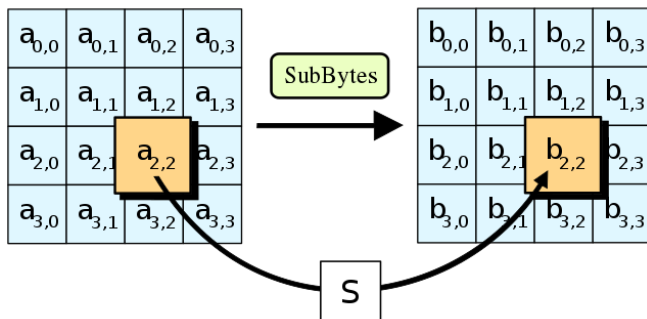
For instance, if there are 16 bytes, these bytes are represented as this matrix:

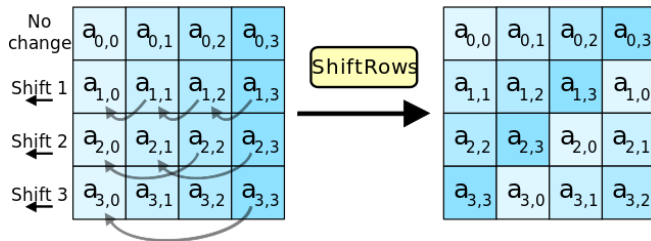
$$\begin{bmatrix} b_0 & b_4 & b_8 & b_{12} \\ b_1 & b_5 & b_9 & b_{13} \\ b_2 & b_6 & b_{10} & b_{14} \\ b_3 & b_7 & b_{11} & b_{15} \end{bmatrix}$$

The key size used for an AES cipher specifies the number of repetitions of transformation rounds. The number of cycles of repetition are as follows:

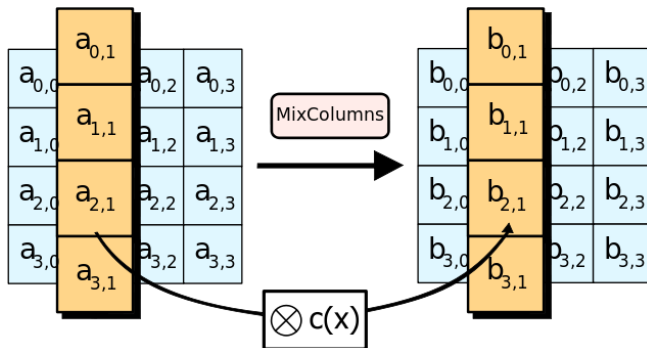
- ▶ 10 cycles of repetition for 128-bit keys.
- ▶ 12 cycles of repetition for 192-bit keys.
- ▶ 14 cycles of repetition for 256-bit keys.

- 1 KeyExpansions – round keys are derived from the cipher key using Rijndael's key schedule. AES requires a separate 128-bit round key block for each round plus one more.
- 2 InitialRound AddRoundKey – each byte of the state is combined with a block of the round key using bitwise xor.
- 3 Rounds
  - ▶ SubBytes – a non-linear substitution step where each byte is replaced with another according to a lookup table.
  - ▶ ShiftRows – a transposition step where the last three rows of the state are shifted cyclically a certain number of steps.
  - ▶ MixColumns – a mixing operation which operates on the columns of the state, combining the four bytes in each column.
  - ▶ AddRoundKey
- 4 Final Round (no MixColumns)
  - ▶ SubBytes
  - ▶ ShiftRows
  - ▶ AddRoundKey.

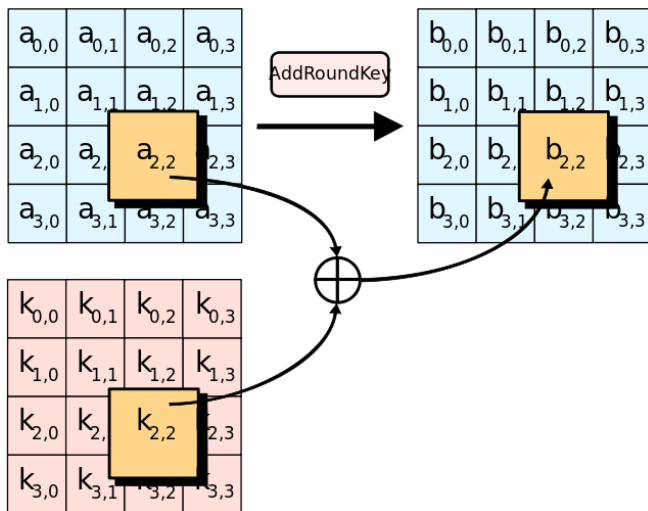




# MixColumns



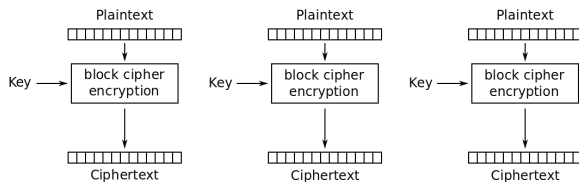
# AddRoundKey



- ▶ Electronic Codebook (ECB)
- ▶ Cipher Block Chaining (CBC)
- ▶ Cipher Feedback (CFB)
- ▶ Output Feedback (OFB)
- ▶ Counter (CTR)

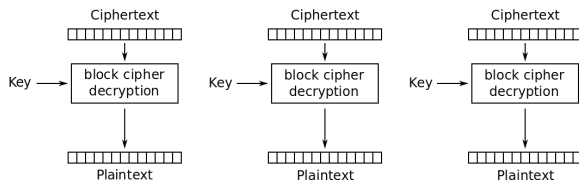


# Electronic Codebook (ECB)



Electronic Codebook (ECB) mode encryption

# Electronic Codebook (ECB)



Electronic Codebook (ECB) mode decryption

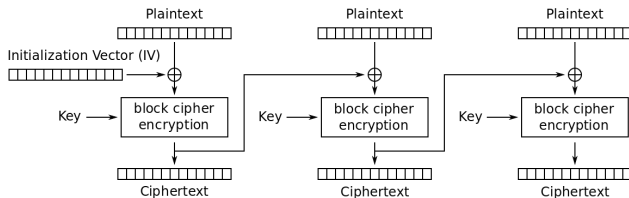
# Electronic Codebook (ECB)



Advantages: parallel encryption

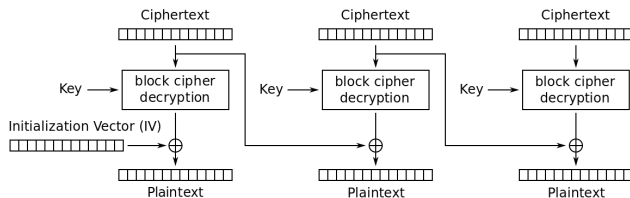
Disadvantages: identical plaintext blocks are encrypted into identical ciphertext blocks; thus, it does not hide data patterns well. In some senses, it doesn't provide serious message confidentiality, and it is not recommended for use in cryptographic protocols at all.

# Cipher Block Chaining (CBC)



Cipher Block Chaining (CBC) mode encryption

# Cipher Block Chaining (CBC)



Cipher Block Chaining (CBC) mode decryption

CBC has been the most commonly used mode of operation. Its main drawbacks are that encryption is sequential (i.e., it cannot be parallelized), and that the message must be padded to a multiple of the cipher block size.

- 1 Introduction to cryptography
- 2 Information-theoretic security
  - Measure of information
  - Perfect secrecy
- 3 Secret key cryptography
- 4 Hash functions

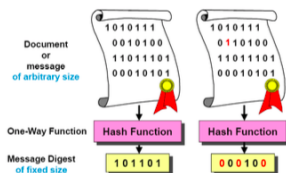


The following mapping  $h : \{0, 1\}^* \rightarrow \{0, 1\}^\ell$  is called a hash function.

# One-way hash function

Let  $h$  be a hash function.  $h$  is called one-way hash function if the following conditions hold:

- Evaluation of  $h(X)$  is a simple operation;
- It is computationally hard task to find  $h^{-1}(Y)$ .



- ▶  $h : \{0, 1\}^* \rightarrow \{0, 1\}^\ell$ 
  - ▶ Maps arbitrary strings to strings of fixed length
- ▶ One-way
- ▶ Weak collision resistance
  - ▶ For a given  $X$  it is difficult to find  $Y \neq X$ , such that  $h(X) = h(Y)$
- ▶ Strong collision resistance
  - ▶ Hard to find any distinct  $X$  and  $Y$ , such that  $h(X) = h(Y)$
- ▶ Security strength of  $h$  only depends on  $\ell$ .
- ▶ Common examples: SHA-1, MD5

$k$ -bit hash function,  $N = 2^k$ . Assume, that the output values have uniform distribution.

The probability, that  $n$  values have different hash is as follows:

$$\hat{P}(n) = \prod_{i=0}^{n-1} (1 - i/N) \leq e^{\frac{n(n-1)}{2N}}.$$

Thus, the probability of collision can be calculated as follows

$$P(n) = 1 - \hat{P}(n)$$

$$n_{0.5} \approx \sqrt{2 \log 2} \cdot 2^{k/2}.$$

$$h_i = f(M_i, h_{i-1}).$$

- ▶ Password files (one-way)
- ▶ Digital signatures – Sign hash of message instead of entire message (collision resistant)
- ▶ Data integrity – Compute and store hash of some data & Check later by re-computing hash and comparing (collision resistant)

Thank you for your attention!