CSCI-GA.3205 Applied Cryptography & Network Security

Department of Computer Science
New York University

PRESENTED BY DR. MAZDAK ZAMANI
mazdak.zamani@NYU.edu

Stream Cipher
Block Cipher

03

NYU

Computational ciphers and semantic security

- As we have seen in Shannon's theorem, the only way to achieve perfect security is to have keys that are as long as messages. However, this is quite impractical.
- The only way around Shannon's theorem is to relax our security requirements.
- The way we shall do this is to consider not all possible adversaries, but only *computationally feasible* adversaries, that is, “real world” adversaries that must perform their calculations on real computers using a reasonable amount of time and memory.

Definition of a computational cipher

- A computational cipher $\xi = (E; D)$ is a pair of efficient algorithms, E and D. The encryption algorithm E takes as input a key k, along with a message m, and produces as output a ciphertext c.
- The decryption algorithm D takes as input a key k, a ciphertext c, and outputs a message m.
- Keys lie in some finite key space K, messages lie in a finite message space M, and ciphertexts lie in some finite ciphertext space C.

Definition of a computational cipher – con'd

- Just as for a Shannon cipher, we say that ξ is defined over (K; M; C).
- We will allow the encryption function E to be a probabilistic algorithm $c \leftarrow \text{Enc}(k, m)$
- From now on, whenever we refer to a cipher, we shall mean a computational cipher, as defined above. Moreover, if the encryption algorithm happens to be deterministic, then we may call the cipher a deterministic cipher.

Definition of semantic security

To motivate the definition of semantic security, consider a deterministic cipher $\xi = (E, D)$, defined over (K, M, C).
For all predicates φ on the ciphertext space, and all messages m_0, m_1, we have

$$\Pr[\varphi(E(k, m_0))] = \Pr[\varphi(E(k, m_1))]$$

Where k is a random variable uniformly distributed over the key space K. Instead of insisting that these probabilities are equal, we shall only require that they are very close; that is,

$$\Pr[\varphi(E(k, m_0))] - \Pr[\varphi(E(k, m_1))] \leq \varepsilon$$

for some very small, or *negligible*, value of ε.

NYU

Example of semantic security

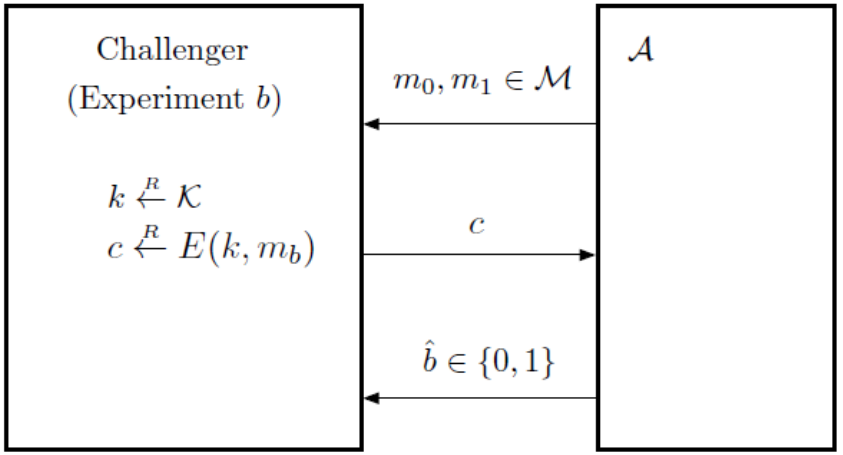
- Suppose it were the case that using the best possible algorithms for generating m_0, and m_1, and for testing some predicate φ, and using (say) 10,000 computers in parallel for 10 years to perform these calculations, it holds for $\varepsilon = 2^{-100}$.
- While not perfectly secure, we might be willing to say that the cipher is *secure for all practical purposes*.

Attack Game

Attack Game 2.1 (semantic security). For a given cipher $\mathcal{E} = (E, D)$, defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$, and for a given adversary $\mathcal{A}$, we define two experiments, Experiment 0 and Experiment 1. For $b = 0, 1$, we define

Experiment b**:**

- The adversary computes $m_0, m_1 \in \mathcal{M}$, of the same length, and sends them to the challenger.
- The challenger computes $k \xleftarrow{R} \mathcal{K}$, $c \xleftarrow{R} E(k, m_b)$, and sends c to the adversary.
- The adversary outputs a bit $\hat{b} \in \{0, 1\}$.



Semantically secure

For $b = 0, 1$, let W_b be the event that $\mathcal{A}$ outputs 1 in Experiment b. We define $\mathcal{A}$'s **semantic security advantage** with respect to $\mathcal{E}$ as

$$\mathsf{SSadv}[\mathcal{A}, \mathcal{E}] := \Big|\Pr[W_0] - \Pr[W_1]\Big|. \quad \square$$

Note that in the above game, the events W_0 and W_1 are defined with respect to the probability space determined by the random choice of k, the random choices made (if any) by the encryption algorithm, and the random choices made (if any) by the adversary. The value $\mathsf{SSadv}[\mathcal{A}, \mathcal{E}]$ is a number between 0 and 1.

Definition 2.2 (semantic security). *A cipher $\mathcal{E}$ is **semantically secure** if for all efficient adversaries $\mathcal{A}$, the value $\mathsf{SSadv}[\mathcal{A}, \mathcal{E}]$ is negligible.*

An *efficient adversary* is one that runs in a "reasonable" amount time.

Stream ciphers

- *Stream cipher* is one type of cipher that allows us to build secure ciphers (semantic security) that use reasonably short keys.

Pseudo-random generators

Recall the one-time pad. Here, keys, messages, and ciphertexts are all L-bit strings. However, we would like to use a key that is much shorter. So the idea is to instead use a short, ℓ-bit "seed" s as the encryption key, where ℓ is much smaller than L, and to "stretch" this seed into a longer, L-bit string that is used to mask the message (and unmask the ciphertext). The string s is stretched using some efficient, deterministic algorithm G that maps ℓ-bit strings to L-bit strings. Thus, the key space for this modified one-time pad is $\{0,1\}^\ell$, while the message and ciphertext spaces are $\{0,1\}^L$. For $s \in \{0,1\}^\ell$ and $m, c \in \{0,1\}^L$, encryption and decryption are defined as follows:

$$E(s,m) := G(s) \oplus m \quad \text{and} \quad D(s,c) := G(s) \oplus c.$$

This modified one-time pad is called a **stream cipher**, and the function G is called a **pseudo-random generator**.

Pseudo-random generators – con'd

- If $l < L$, then by Shannon's Theorem, this stream cipher cannot achieve perfect security; however, if G satisfies an appropriate security property, then this cipher is semantically secure.
- Suppose s is a random l-bit string and r is a random L-bit string.
- Intuitively, if an adversary cannot effectively tell the difference between G(s) and r, then he should not be able to tell the difference between this stream cipher and a one-time pad.

Statistical test

- An algorithm that is used to distinguish a pseudo-random string G(s) from a truly random string r is called a statistical test.
- It takes a string as input, and outputs 0 or 1.
- Such a test is called effective if the probability that it outputs 1 on a pseudo-random input is significantly different than the probability that it outputs 1 on a truly random input.

Designing an effective statistical test

- Given an L-bit string, calculate some statistic, and then see if this statistic differs greatly from what one would expect if the string were truly random.

Example of an effective statistical test

For example, a very simple statistic that is easy to compute is the number k of 1's appearing in the string. For a truly random string, we would expect $k \approx L/2$. If the PRG G had some bias towards either 0-bits or 1-bits, we could effectively detect this with a statistical test that, say, outputs 1 if $|k - 0.5L| < 0.01L$, and otherwise outputs 0. This statistical test would be quite effective if the PRG G did indeed have some significant bias towards either 0 or 1.

Definition of a pseudo-random generator

A **pseudo-random generator**, or **PRG** for short, is an efficient, deterministic algorithm G that, given as input a **seed** s, computes an output r. The seed s comes from a finite **seed space** $\mathcal{S}$ and the output r belongs to a finite **output space** $\mathcal{R}$. Typically, $\mathcal{S}$ and $\mathcal{R}$ are sets of bit strings of some prescribed length (for example, in the discussion above, we had $\mathcal{S} = \{0,1\}^\ell$ and $\mathcal{R} = \{0,1\}^L$). We say that G is a PRG defined over $(\mathcal{S}, \mathcal{R})$.

Our definition of security for a PRG captures the intuitive notion that if s is chosen at random from $\mathcal{S}$ and r is chosen at random from $\mathcal{R}$, then no efficient adversary can effectively tell the difference between $G(s)$ and r: the two are **computationally indistinguishable**. The definition is formulated as an attack game.

Attack Game

Attack Game 3.1 (PRG). For a given PRG G, defined over $(\mathcal{S}, \mathcal{R})$, and for a given adversary $\mathcal{A}$, we define two experiments, Experiment 0 and Experiment 1. For $b = 0, 1$, we define:

Experiment b**:**

- if $b = 0$: $s \xleftarrow{R} \mathcal{S}$, $r \leftarrow G(s)$;
- if $b = 1$: $r \xleftarrow{R} \mathcal{R}$.

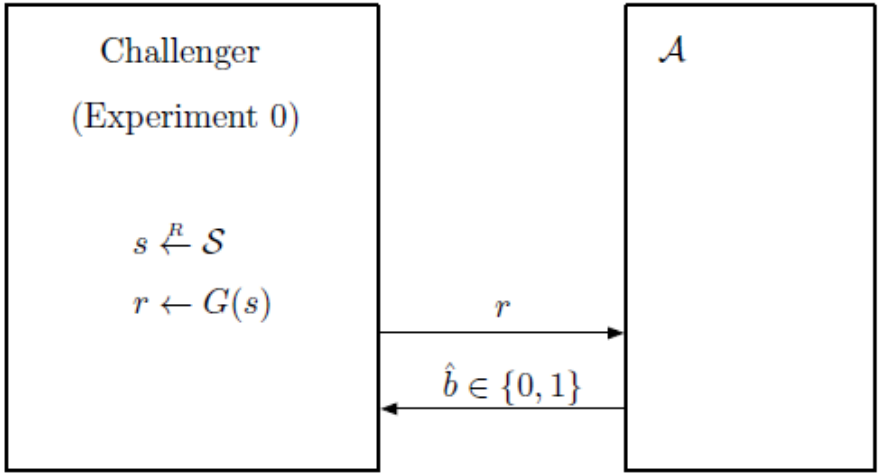
and sends r to the adversary.

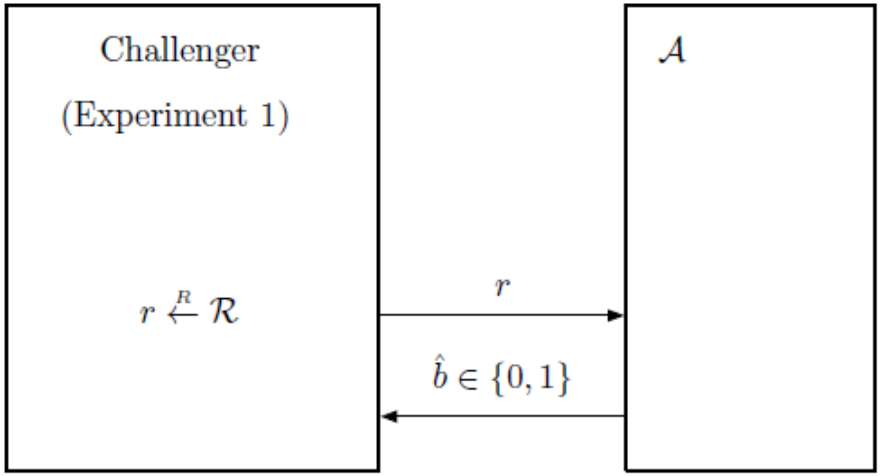
- Given r, the adversary computes and outputs a bit $\hat{b} \in \{0, 1\}$.

For $b = 0, 1$, let W_b be the event that $\mathcal{A}$ outputs 1 in Experiment b. We define $\mathcal{A}$'s **advantage** with respect to G as

$$\mathsf{PRGadv}[\mathcal{A}, G] := \Big|\Pr[W_0] - \Pr[W_1]\Big|. \quad \square$$

Experiments 0 and 1 of Attack Game 3.1





Stream ciphers: encryption with a PRG

Let G be a PRG defined over $(\{0,1\}^\ell, \{0,1\}^L)$; that is, G stretches an ℓ-bit seed to an L-bit output. The **stream cipher $\mathcal{E} = (E, D)$ constructed from** G is defined over $(\{0,1\}^\ell, \{0,1\}^{\leq L}, \{0,1\}^{\leq L})$; for $s \in \{0,1\}^\ell$ and $m, c \in \{0,1\}^{\leq L}$, encryption and decryption are defined as follows: if $|m| = v$, then

$$E(s, m) := G(s)[0\,..\,v-1] \;\oplus\; m,$$

and if $|c| = v$, then

$$D(s, c) := G(s)[0\,..\,v-1] \;\oplus\; c.$$

Stream cipher limitations: attacks on the one-time pad

- Although stream ciphers are semantically secure, they are highly brittle and become totally insecure if used incorrectly.
  - The two-time pad is insecure
  - The one-time pad is malleable

The two-time pad is insecure

A stream cipher is well equipped to encrypt a *single* message from Alice to Bob. Alice, however, may wish to send several messages to Bob. For simplicity suppose Alice wishes to encrypt two messages m_1 and m_2. The naive solution is to encrypt both messages using the same stream cipher key s:

$$c_1 \leftarrow m_1 \oplus G(s) \qquad \text{and} \qquad c_2 \leftarrow m_2 \oplus G(s) \tag{3.8}$$

A moments reflection shows that this construction is insecure in a very strong sense. An adversary who intercepts c_1 and c_2 can compute

$$\Delta := c_1 \oplus c_2 = \big(m_1 \oplus G(s)\big) \oplus \big(m_2 \oplus G(s)\big) = m_1 \oplus m_2$$

and obtain the xor of m_1 and m_2. Not surprisingly, English text contains enough redundancy that given $\Delta = m_1 \oplus m_2$ the adversary can recover both m_1 and m_2 in the clear. Hence, the construction in (3.8) leaks the plaintexts after seeing only two sufficiently long ciphertexts.

The one-time pad is malleable

Although semantic security ensures that an adversary cannot read the plaintext, it provides no guarantees for integrity. When using a stream cipher, an adversary can change a ciphertext and the modification will never be detected by the decryptor. Even worse, let us show that by changing the ciphertext, the attacker can control how the decrypted plaintext will change.

Suppose an attacker intercepts a ciphertext $c := E(s, m) = m \oplus G(s)$. The attacker changes c to $c' := c \oplus \Delta$ for some Δ of the attacker's choice. Consequently, the decryptor receives the modified message

$$D(s, c') = c' \oplus G(s) = (c \oplus \Delta) \oplus G(s) = m \oplus \Delta.$$

Linear-Feedback Shift Registers

- The linear-feedback shift registers (LFSRs) have been used historically for pseudorandom-number generation, as they are extremely efficient to implement in hardware, and generate output having good statistical properties.
- By themselves, however, they do not give cryptographically strong pseudorandom generators, and in fact we will show an easy key recovery attack on LFSRs. Nevertheless, LFSRs can be used as a component in building stream ciphers with better security.

A linear-feedback shift register

Linear feedback shift registers (LFSR). The CSS stream cipher is built from two LFSRs. An n-bit LFSR is defined by a set of integers $V := \{v_1, \ldots, v_d\}$ where each v_i is in the range $\{0, \ldots, n-1\}$. The elements of V are called **tap positions**. An LFSR gives a PRG as follows (Fig. 3.10):

Input: $s = (b_{n-1}, \ldots, b_0) \in \{0,1\}^n$ and $s \neq 0^n$
Output: $y \in \{0,1\}^\ell$ where $\ell > n$

for $i \leftarrow 1 \ldots \ell$ do
    output b_0 // *output one bit*
    $b \leftarrow b_{v_1} \oplus \cdots \oplus b_{v_d}$ // *compute feedback bit*
    $s \leftarrow (b,\ b_{n-1}, \ldots,\ b_1)$ // *shift register bits to the right*

NYU

A linear-feedback shift register

The LFSR outputs one bit per clock cycle. Note that if an LFSR is started in state s = 0^n then its output is degenerate, namely all 0. For this reason one of the seed bits is always set to 1.

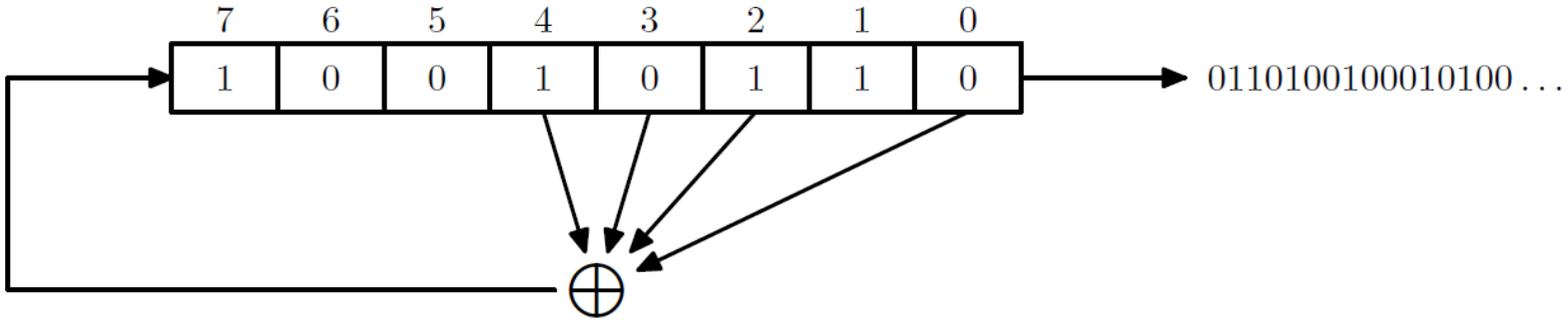


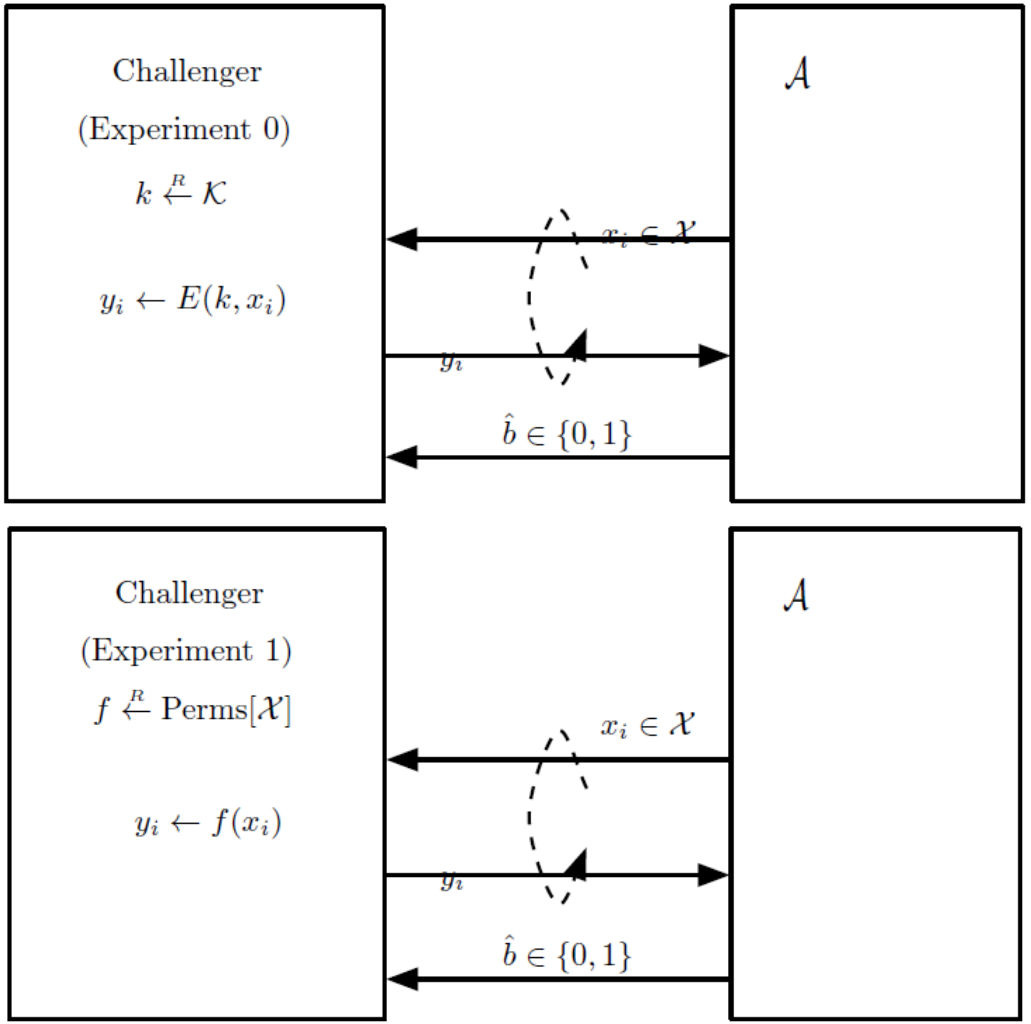
Figure 3.10: The 8 bit linear feedback shift register $\{4, 3, 2, 0\}$

Block ciphers

Functionally, a **block cipher** is a deterministic cipher $\mathcal{E} = (E, D)$ whose message space and ciphertext space are the same (finite) set $\mathcal{X}$. If the key space of $\mathcal{E}$ is $\mathcal{K}$, we say that $\mathcal{E}$ is a block cipher **defined over** $(\mathcal{K}, \mathcal{X})$. We call an element $x \in \mathcal{X}$ a **data block**, and refer to $\mathcal{X}$ as the **data block space** of $\mathcal{E}$.

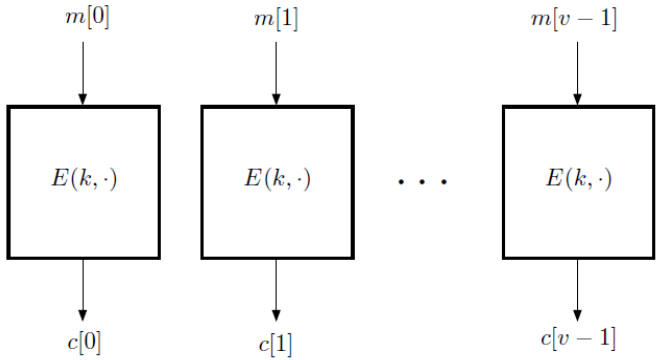
For every fixed key $k \in \mathcal{K}$, we can define the function $f_k := E(k, \cdot)$; that is, $f_k : \mathcal{X} \to \mathcal{X}$ sends $x \in \mathcal{X}$ to $E(k, x) \in \mathcal{X}$. The usual correctness requirement for any cipher implies that for every fixed key k, the function f_k is one-to-one, and as $\mathcal{X}$ is finite, f_k must be onto as well. Thus, f_k is a permutation on $\mathcal{X}$, and $D(k, \cdot)$ is the inverse permutation f_k^{-1}.

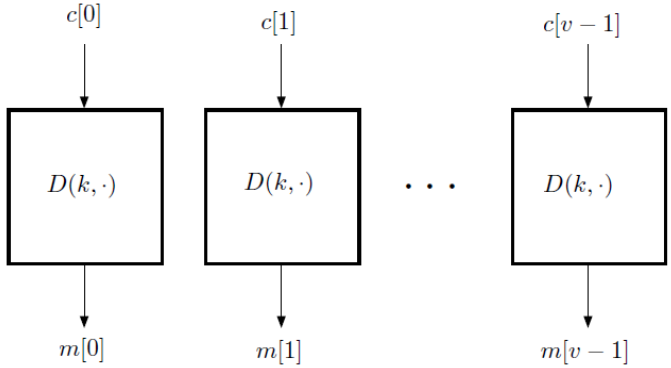
A secure block cipher is unpredictable



Electronic codebook mode

- If we want to encrypt longer messages, a natural idea would be to break up a long message into a sequence of data blocks and encrypt each data block separately.
- This use of a block cipher to encrypt long messages is called electronic codebook mode, or ECB mode for short.





Pseudo Random Permutation

- Pseudo Random Function (**PRF**) defined over (K,X,Y):

  $F: K \times X \rightarrow Y$

  such that exists "efficient" algorithm to evaluate F(k,x)
- Pseudo Random Permutation (**PRP**) defined over (K,X):

  $E: K \times X \rightarrow X$

such that:

1. Exists "efficient" deterministic algorithm to evaluate E(k,x)
2. The function $E(k, \cdot)$ is one-to-one
3. Exists "efficient" inversion algorithm D(k,y)

Constructing block ciphers in practice

Block ciphers are a basic primitive in cryptography from which many other systems are built. Virtually all block ciphers used in practice use the same basic framework called the **iterated cipher** paradigm. To construct an iterated block cipher the designer makes two choices:

- First, he picks a simple block cipher $\hat{\mathcal{E}} := (\hat{E}, \hat{D})$ that is clearly insecure on its own. We call $\hat{\mathcal{E}}$ the **round cipher**.
- Second, he picks a simple (not necessarily secure) PRG G that is used to expand the key k into d keys $k_1, \ldots, k_d$ for $\hat{\mathcal{E}}$. We call G the **key expansion function**.

NYU

Algorithm $E(k, x)$:

- step 1. **key expansion**: use the key expansion function G to stretch the key k of $\mathcal{E}$ to d keys of $\hat{\mathcal{E}}$:

$$(k_1, \ldots, k_d) \leftarrow G(k)$$

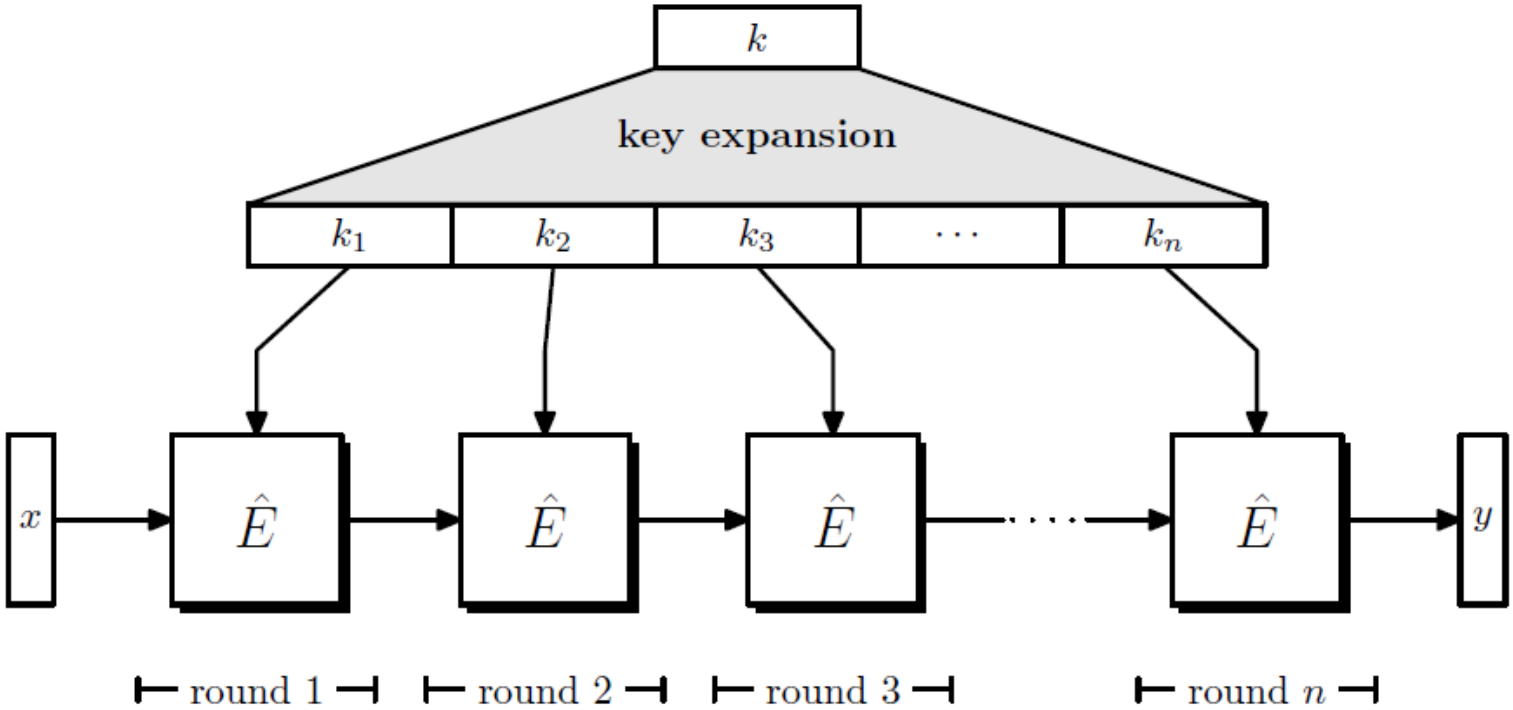
- step 2. **iteration**: for $i = 1, \ldots, d$ apply $\hat{E}(k_i, \cdot)$, namely:

$$y \leftarrow \hat{E}(k_d,\ \hat{E}(k_{d-1}, \ldots, \hat{E}(k_2,\ \hat{E}(k_1,\ x)) \ldots))$$

Each application of $\hat{E}$ is called a **round** and the total number of rounds is d. The keys $k_1, \ldots, k_d$ are called **round keys**. The decryption algorithm $D(k, y)$ is identical except that the round keys are applied in reverse order. $D(k, y)$ is defined as:

$$x \leftarrow \hat{D}(k_1,\ \hat{D}(k_2, \ldots, \hat{D}(k_{d-1},\ \hat{D}(k_d,\ y)) \ldots))$$

Block Ciphers Built by Iteration



Sample block ciphers

	key size (bits)	block size (bits)	number of rounds	performance[1] (MB/sec)
DES	56	64	16	80
3DES	168	64	48	30
AES-128	128	128	10	163
AES-256	256	128	14	115

The DES algorithm

The DES algorithm consists of 16 iterations of a simple round cipher. To describe DES it suffices to describe the DES round cipher and the DES key expansion function. We describe each in turn.

The Feistel permutation. One of the key innovations in DES, invented by Horst Feistel at IBM, builds a permutation from an arbitrary function. Let $f : \mathcal{X} \to \mathcal{X}$ be a function. We construct a permutations $\pi : \mathcal{X}^2 \to \mathcal{X}^2$ as follows (Fig. 4.7):

$$\pi(x, y) := \big(y,\ x \oplus f(y)\big)$$

To show that π is one-to-one we construct its inverse, which is given by:

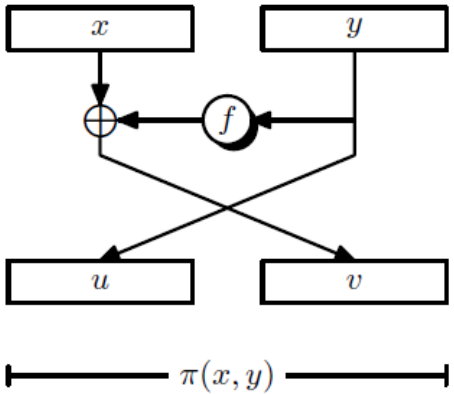
$$\pi^{-1}(u, v) = \big(v \oplus f(u),\ u\big)$$

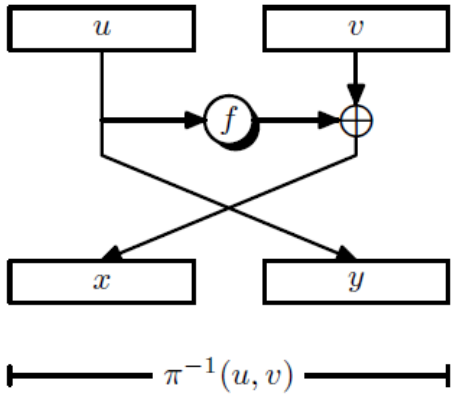
NYU

Feistel permutation

The function π is called a **Feistel permutation** and is used to build the DES round cipher. The composition of n Feistel permutations is called an **n-round Feistel network**. Block ciphers designed as a Feistel network are called **Feistel ciphers**. For DES, the function f takes 32-bit inputs and the resulting permutation π operates on 64-bit blocks.

Note that the Feistel inverse function π^{-1} is almost identical to π. As a result the same hardware can be used for evaluating both π and π^{-1}. This in turn means that the encryption and decryption circuits can use the same hardware.





NYU

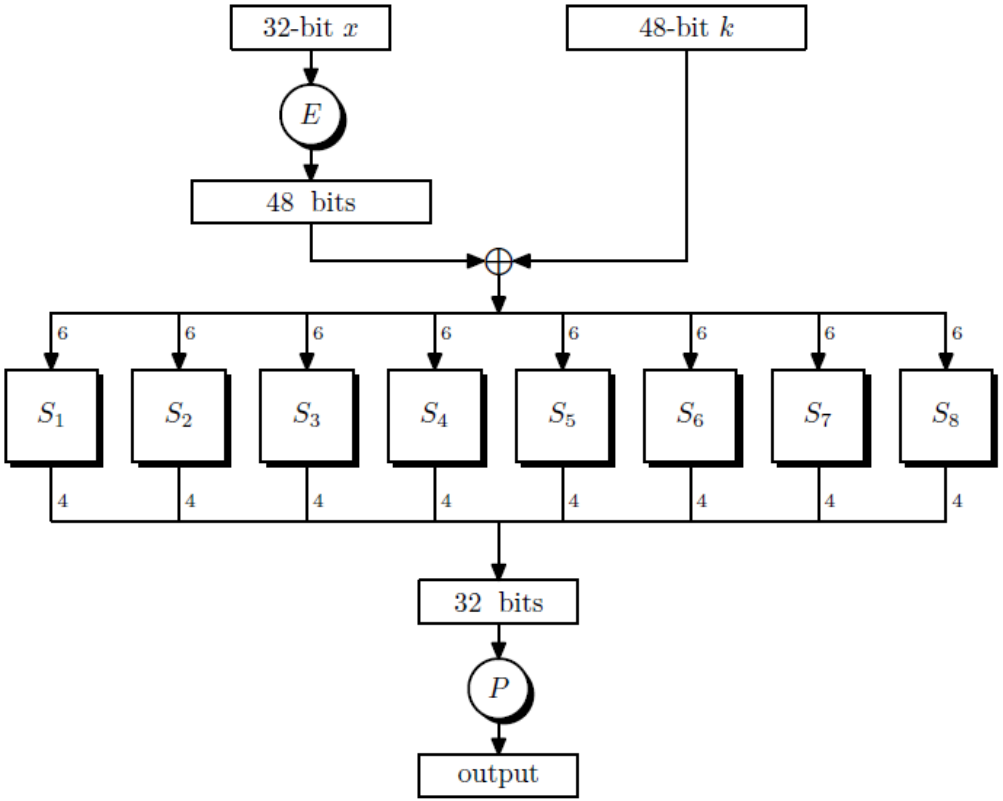
The DES round function $F(k, x)$. The DES encryption algorithm is a 16-round Feistel network where each round uses a different function $f : \mathcal{X} \to \mathcal{X}$. In round number i the function f is defined as

$$f(x) := F(k_i, x)$$

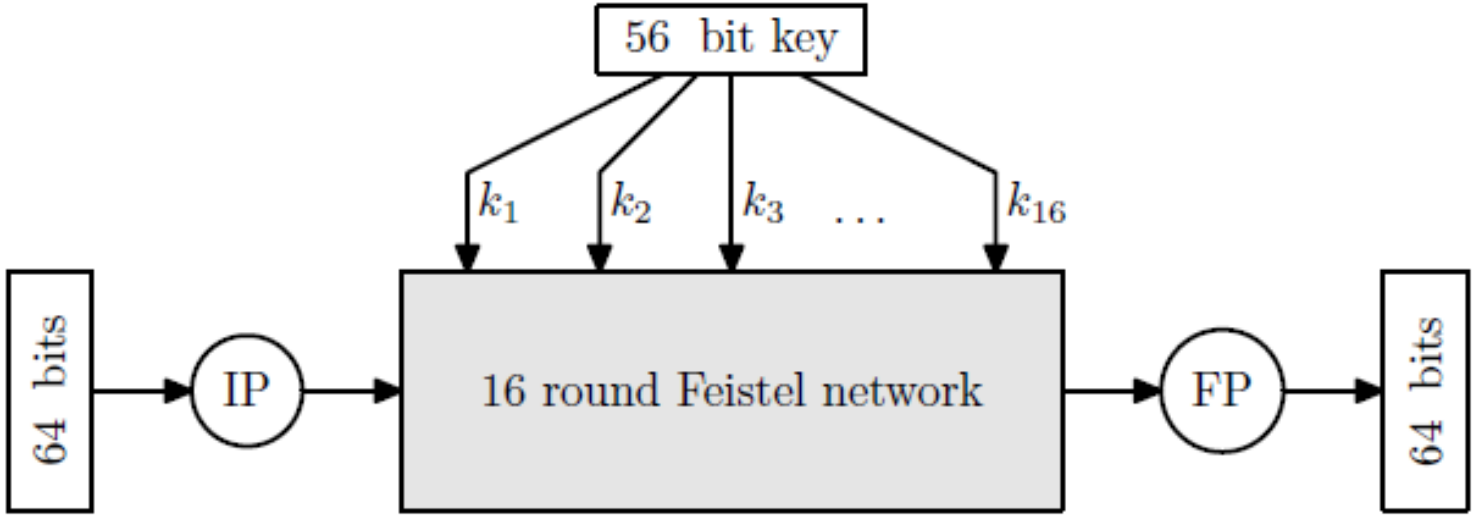
where k_i is a 48-bit key for round number i and F is a fixed function called the **DES round function**. The function F is the centerpiece of the DES algorithm and is shown in Fig. 4.8. F uses several auxiliary functions E, P, and $S_1, \ldots, S_8$ defined as follows:

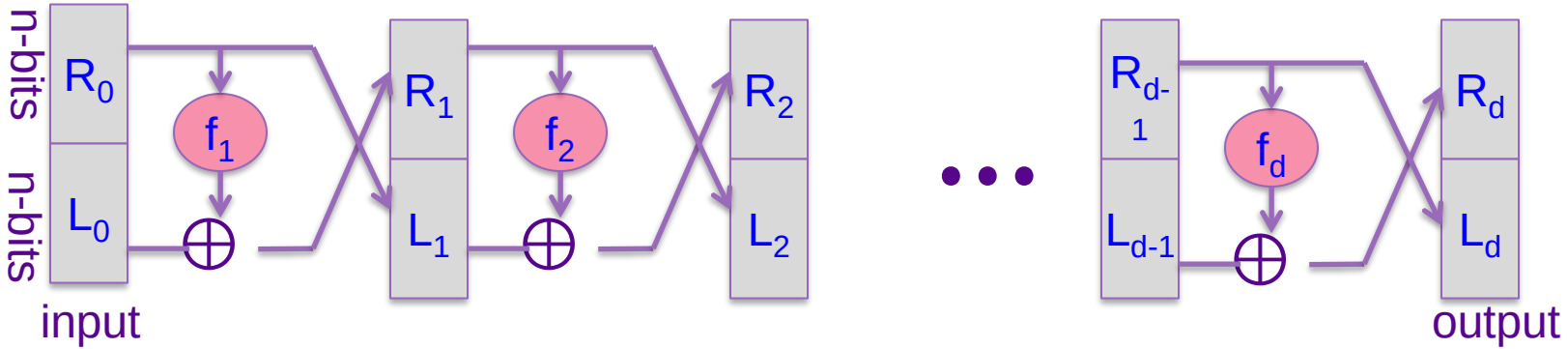
- The function E expands a 32-bit input to a 48-bit output by rearranging and replicating the input bits. For example, E maps input bit number 1 to output bits 2 and 48; it maps input bit 2 to output bit number 3, and so on.
- The function P, called the **mixing permutation**, maps a 32-bit input to a 32-bit output by rearranging the bits of the input. For example, P maps input bit number 1 to output bit number 9; input bit number 2 to output number 15, and so on.
- At the heart of the DES algorithm are the functions $S_1, \ldots, S_8$ called **S-boxes**. Each S-box S_i maps a 6-bit input to a 4-bit output by a lookup table. The DES standard lists these 8 look-up tables, where each table contains 64 entries.

The DES round function F(k; x)



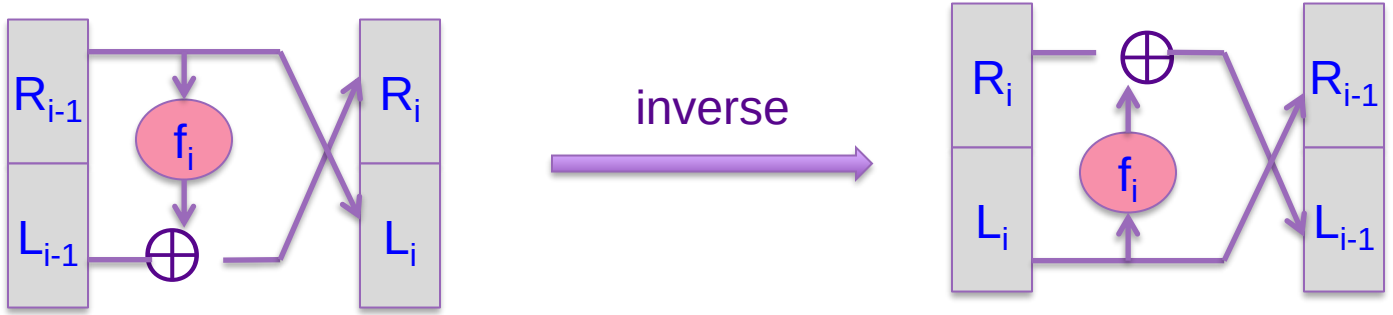
The complete DES circuit





For all $f_1, \ldots, f_d$: $\{0,1\}^n \longrightarrow \{0,1\}^n$

Feistel network $F: \{0,1\}^{2n} \longrightarrow \{0,1\}^{2n}$ is invertible



The S-boxes

- S_i: $\{0,1\}^6 \longrightarrow \{0,1\}^4$

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

NYU

The AES block cipher

- Like many real-world block ciphers, AES is an iterated cipher that iterates a simple round cipher several times. The number of iterations depends on the size of the secret key:

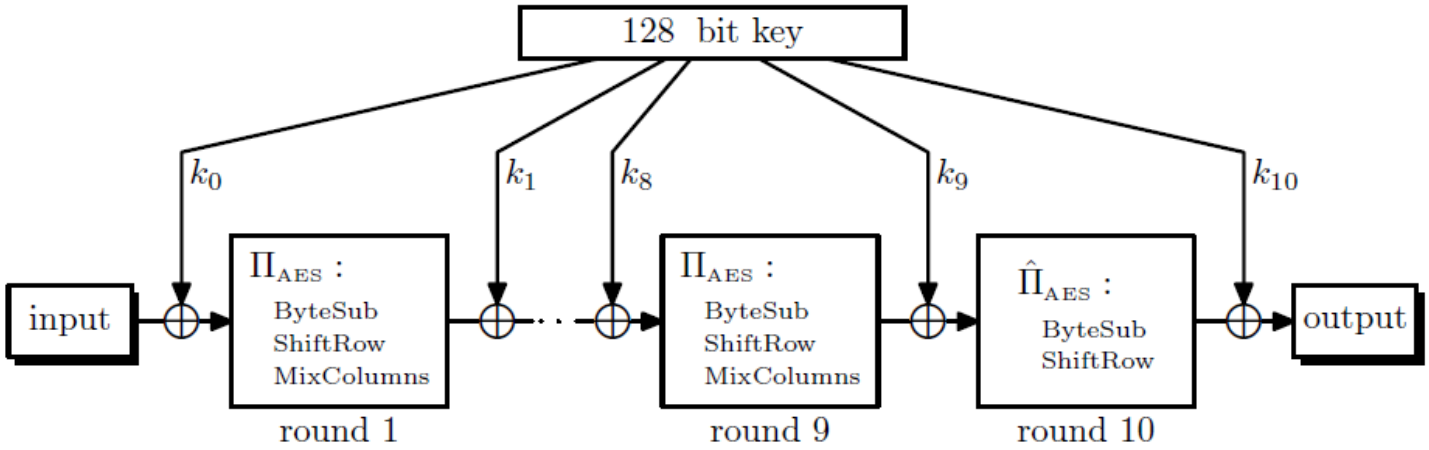
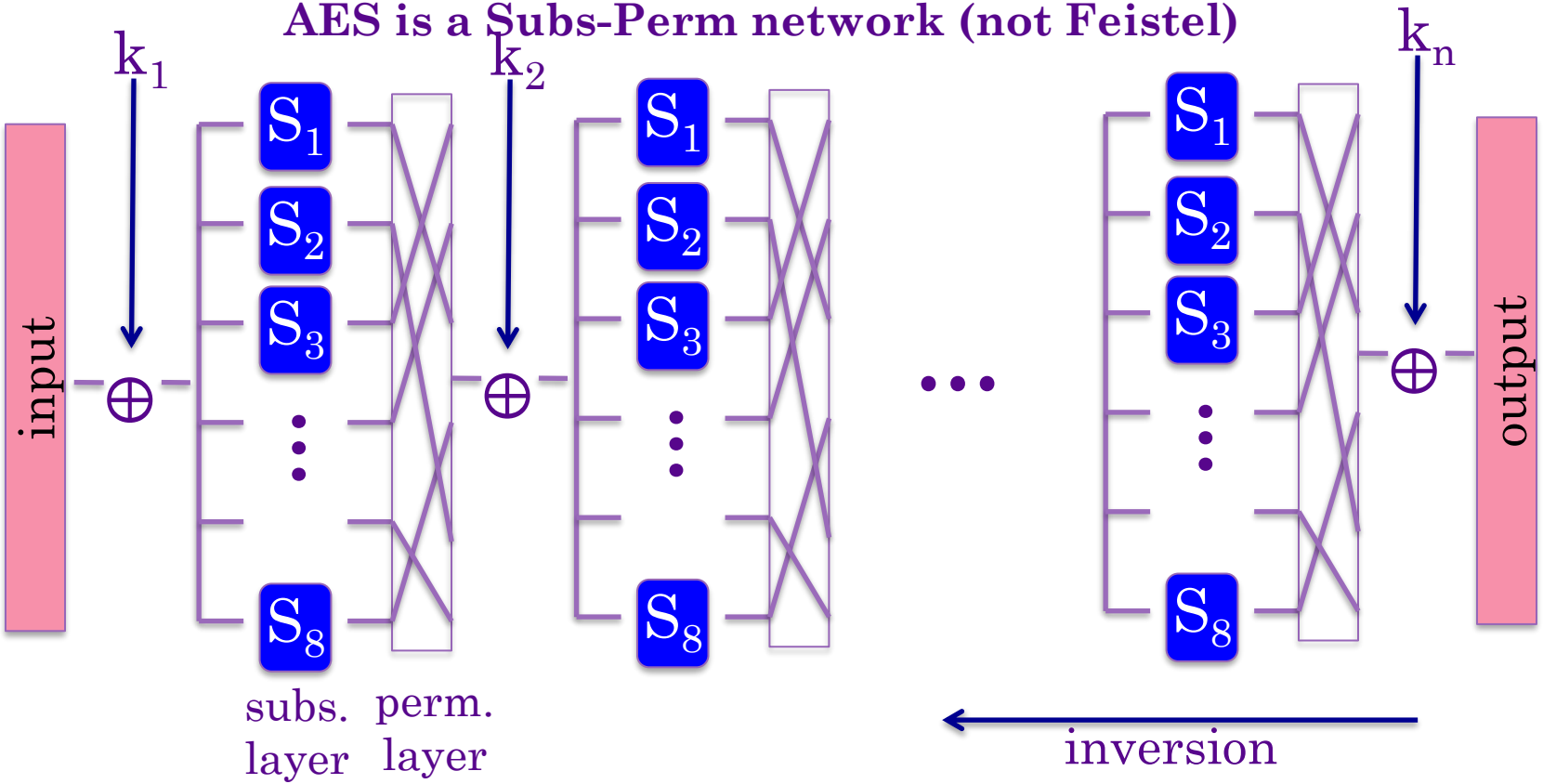
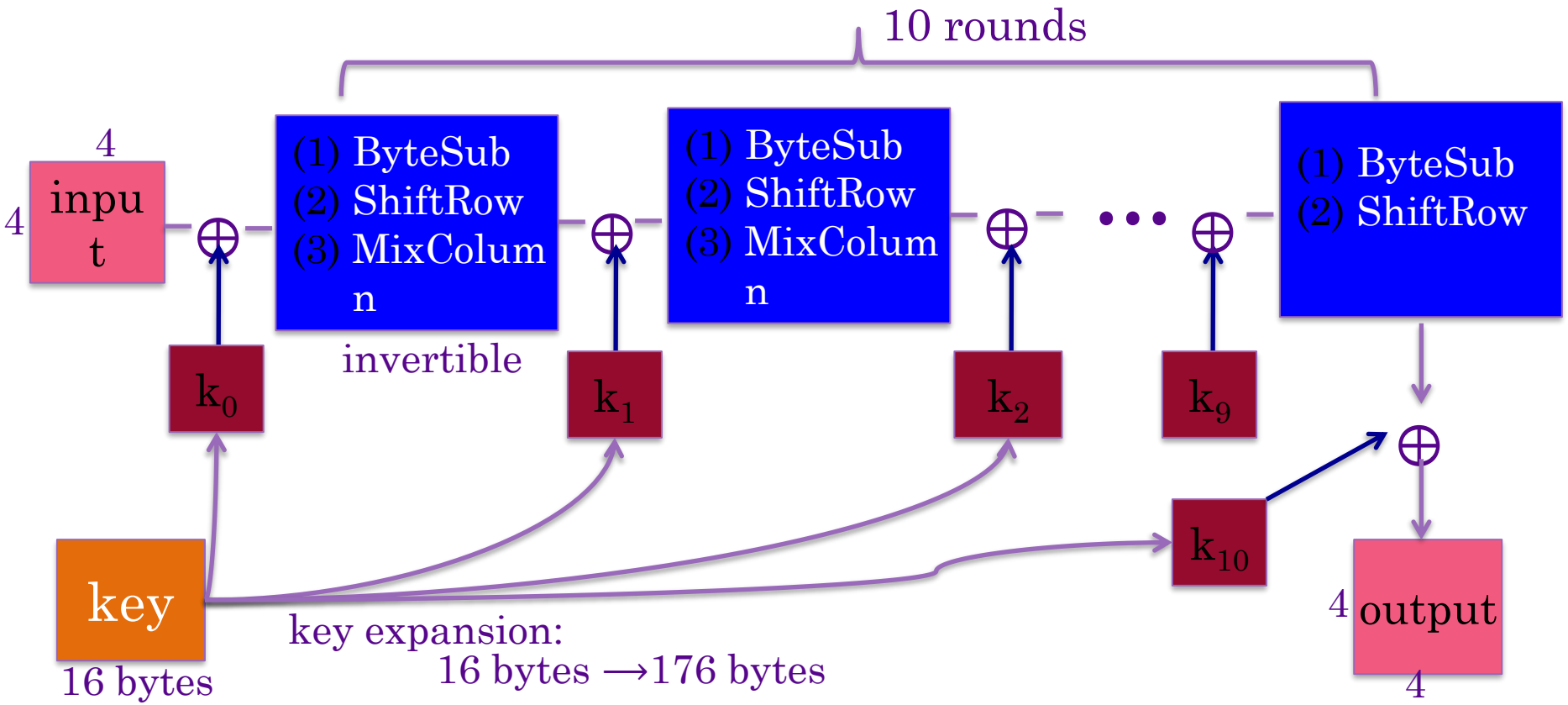


Figure 4.11: Schematic of the AES-128 block cipher

AES is a Subs-Perm network (not Feistel)

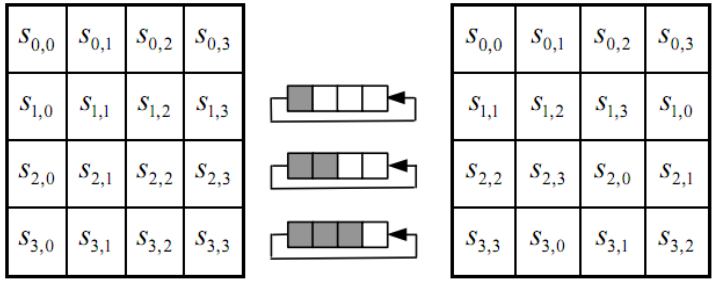


AES-128 schematic

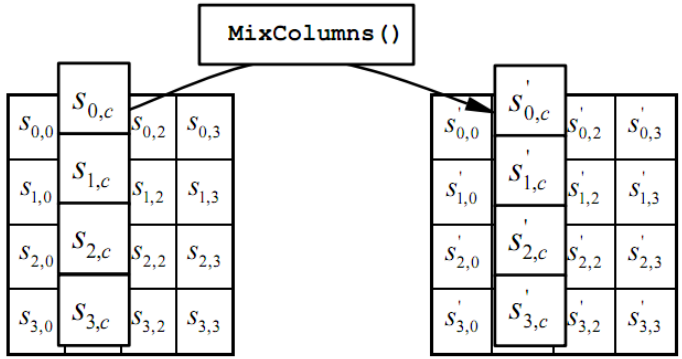


The round function

- **ByteSub**: a 1 byte S-box. 256 byte table (easily computable)
- **ShiftRows**:



- **MixColumns**:



AES Key Sizes

cipher name	key-size (bits)	block-size (bits)	number of rounds
AES-128	128	128	10
AES-192	192	128	12
AES-256	256	128	14

NYU

The AES round permutation

The AES round permutation. The permutation Π_{AES} is made up of a sequence of three invertible operations on the set $\{0,1\}^{128}$. The input 128-bits is organized as a 4×4 array of cells, where each cell is eight bits. The following three invertible operations are then carried out in sequence, one after the other, on this 4×4 array:

1. `SubBytes`: Let $S : \{0,1\}^8 \to \{0,1\}^8$ be a fixed permutation (a one-to-one function). This permutation is applied to each of the 16 cells, one cell at a time. The permutation S is specified in the AES standard as a hard-coded table of 256 entries. It is designed to have no fixed points, namely $S(x) \neq x$ for all $x \in \{0,1\}^8$, and no inverse fixed points, namely $S(x) \neq \bar{x}$ where $\bar{x}$ is the bit-wise complement of x. These requirements are needed to defeat certain attacks discussed in Section 4.3.1.

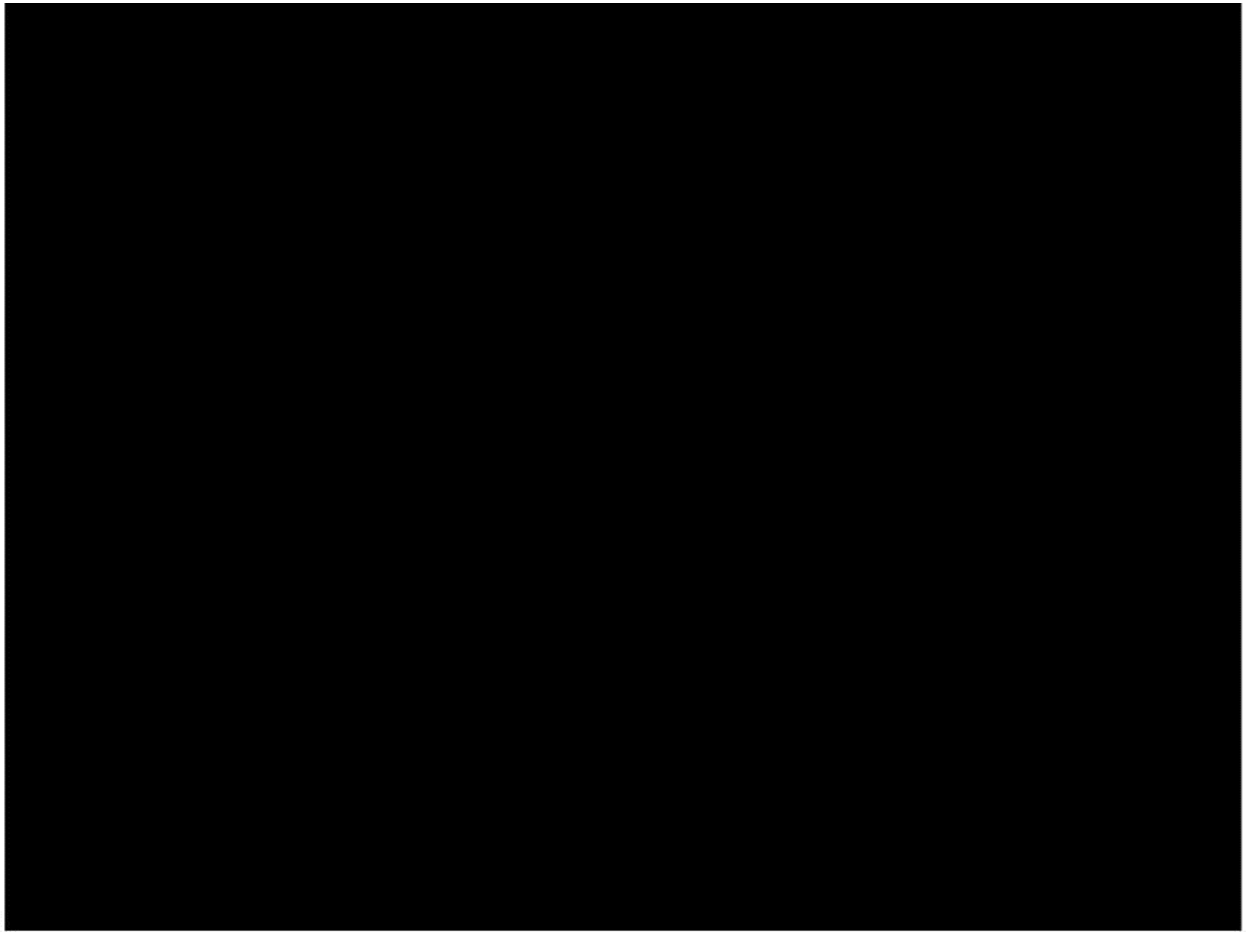
2. `ShiftRows`: This step performs a cyclic shift on the four rows of the input 4×4 array: the first row is unchanged, the second row is cyclically shifted one byte to the left, the third row is cyclically shifted two bytes, and the fourth row is cyclically shifted three bytes. In a diagram, this step performs the following transformation:

$$\begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 & a_7 \\ a_8 & a_9 & a_{10} & a_{11} \\ a_{12} & a_{13} & a_{14} & a_{15} \end{pmatrix} \Longrightarrow \begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ a_5 & a_6 & a_7 & a_4 \\ a_{10} & a_{11} & a_8 & a_9 \\ a_{15} & a_{12} & a_{13} & a_{14} \end{pmatrix} \tag{4.11}$$

3. `MixColumns`: In this step the 4×4 array is treated as a matrix and this matrix is multiplied by a fixed matrix where arithmetic is interpreted in the finite field $\mathrm{GF}(2^8)$. Elements in the field $\mathrm{GF}(2^8)$ are represented as polynomials over $\mathrm{GF}(2)$ of degree less than eight where multiplication is done modulo the irreducible polynomial $x^8 + x^4 + x^3 + x + 1$. Specifically, the `MixColumns` transformation does:

$$\begin{pmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{pmatrix} \times \begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ a_5 & a_6 & a_7 & a_4 \\ a_{10} & a_{11} & a_8 & a_9 \\ a_{15} & a_{12} & a_{13} & a_{14} \end{pmatrix} \Longrightarrow \begin{pmatrix} a'_0 & a'_1 & a'_2 & a'_3 \\ a'_4 & a'_5 & a'_6 & a'_7 \\ a'_8 & a'_9 & a'_{10} & a'_{11} \\ a'_{12} & a'_{13} & a'_{14} & a'_{15} \end{pmatrix} \tag{4.12}$$

Here the scalars $01, 02, 03$ are interpreted as elements of $\mathrm{GF}(2^8)$ using their binary representation (e.g., 03 represents the element $x + 1$ in $\mathrm{GF}(2^8)$). This fixed matrix is invertible over $\mathrm{GF}(2^8)$ so that the entire transformation is invertible.



NYU