Cryptography

Lecture 5

Vahid Amin-Ghafari

Vahidaming@ustc.edu.cn

ENTROPY

- for any probability distribution, we define a quantity called the *entropy*, which has many properties that agree with the intuitive notion of what a measure of information should be
- *Entropy*: is a measure of the uncertainty (information) of a random variable.

آنتروپی: معیاری عددی برای اندازه گرفتن اطلاعات، یا ابهام یک متغیر تصادفی است. به بیان دقیق‌تر، آنتروپی یک متغیر تصادفی، متوسط اطلاعات آن است.

- Let X be a discrete random variable with alphabet $\mathcal{x}$ and probability mass function p(x) = Pr{X = x}, x∈$\mathcal{x}$.
- **Definition**: The entropy $H(X)$ of a discrete random variable X is defined by (The log is to the base 2 and entropy is expressed in bits, **video**)

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x)$$

ENTROPY

تعریف میزان اطلاعات

در این درس می‌خواهیم نحوهٔ اندازه گیری مقدار اطلاعات یک متغیر تصادفی را بیان کنیم.

ابتدا خلاصه ای راجع به اطلاعات موجود در یک متغیر تصادفی بیان می کنیم.

هر متغیر تصادفی دارای مقدار اطلاعات مشخصی می‌باشد.

مقدار اطلاعات موجود در یک متغیر تصادفی برابر مقدار تردیدی است که با روشن شدن آن از ذهن خارج می گردد.

ENTROPY

تعریف میزان اطلاعات – مثال

در یک مسابقهٔ اسب دوانی شماره اسب برنده را به عنوان یک متغیر تصادفی X در نظر می گیریم،

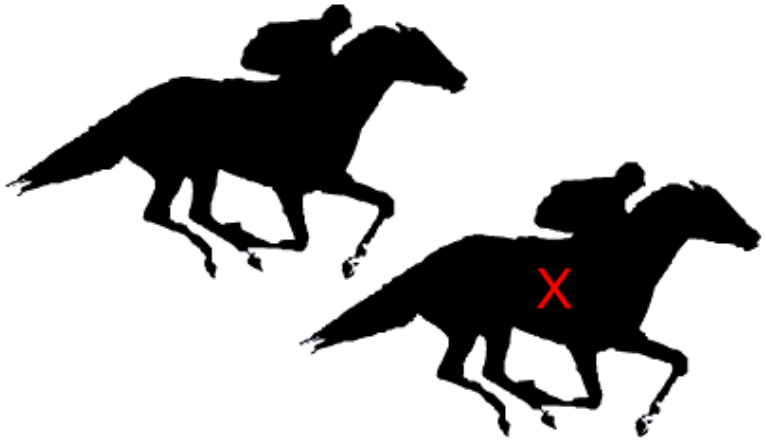
هدف ما اندازه گیری میزان اطلاعات متغیر X است.

با مشخص شدن نتیجهٔ مسابقه تردید یا ابهامی که راجع به متغیر تصادفی X در ذهن ما بود خارج می‌شود،

این ابهام به تعداد اسب های شرکت کننده بستگی دارد.

ENTROPY

اگر تنها دو اسب در مسابقه شرکت نمایند ابهام کمی در مورد نتیجه مسابقه داریم



ENTROPY

در حالیکه اگر ۲۰ اسب در مسابقه شرکت کرده باشند

ابهام بیشتری داریم.

ENTROPY

با فرض مساوی بودن شانس اسب ها، هر چه تعداد اسب های شرکت کننده بیشتر شوند ابهام ما نیز بیشتر می‌شود.

تعریف آنتروپی

ما دو واژهٔ "اطلاعات" موجود در یک متغیر تصادفی X و "تردید" موجود در آن متغیر را معادل یکدیگر به کار می‌بریم و به آن آنتروپی (Entropy) متغیر تصادفی X می‌گوییم.

ENTROPY

تردید راجع به نتیجه مسابقه اسب دوانی علاوه بر تعداد اسب ها به چه چیز دیگری بستگی دارد؟

مسلماً به احتمال برنده شدن اسب ها.

یعنی اگر از قبل بدانیم یکی از اسب ها شانس بسیار زیادی برای برنده شدن دارد

تردید کمی راجع به نتیجه مسابقه داریم

در حالیکه اگر همهٔ اسب ها شانس یکسانی داشته باشند تردیدمان بیشترخواهد بود.

اندازه گیری مقدار آنتروپی یک متغیر تصادفی

آنتروپی یک متغیر تصادفی X که به آن منبع X نیز اطلاق می‌شود

چگونه اندازه گیری می‌شود؟

ENTROPY

- Lemma : H(X) ≥ 0

  Proof**:** 0 ≤ *p(x)* ≤ 1 implies that $\log p(x)$ 1 ≥ 0

- ***Example***

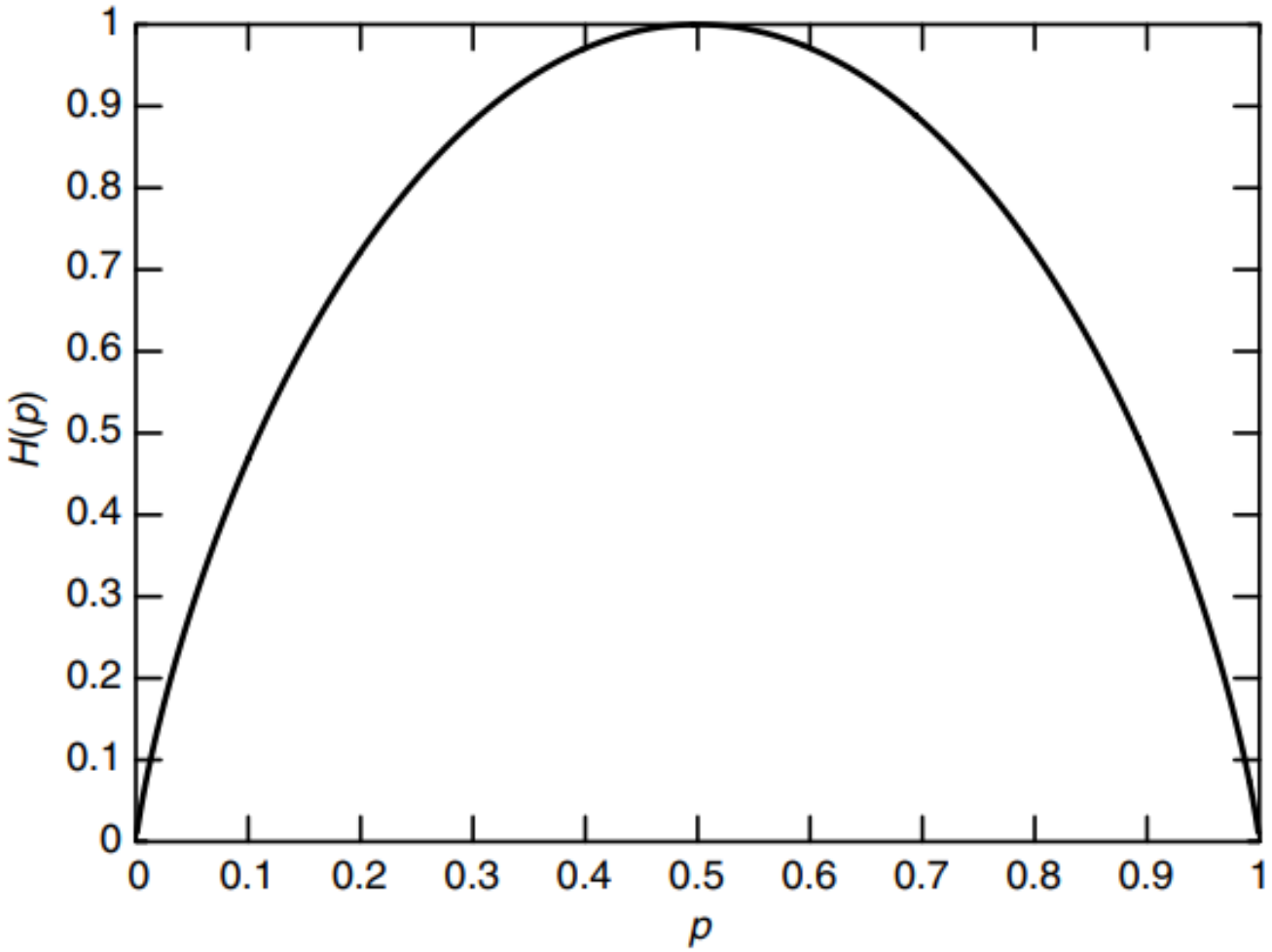
$$X = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } 1 - p \end{cases}$$

- In particular, $H(X) = 1$ bit when $p = {}^1/_2$

$$H(X) = -p \log p - (1-p) \log(1-p) \stackrel{\text{def}}{=\!=} H(p)$$

ENTROPY

- graph of the function $H(p)$



ENTROPY

- It is a concave function of the distribution
- Equals 0 when $p = 0$ or 1. This makes sense, because the variable is not random and there is no uncertainty.
- Uncertainty is maximum when $p = {}^1/_2$, which also corresponds to the maximum value of the entropy.
- ***Example***

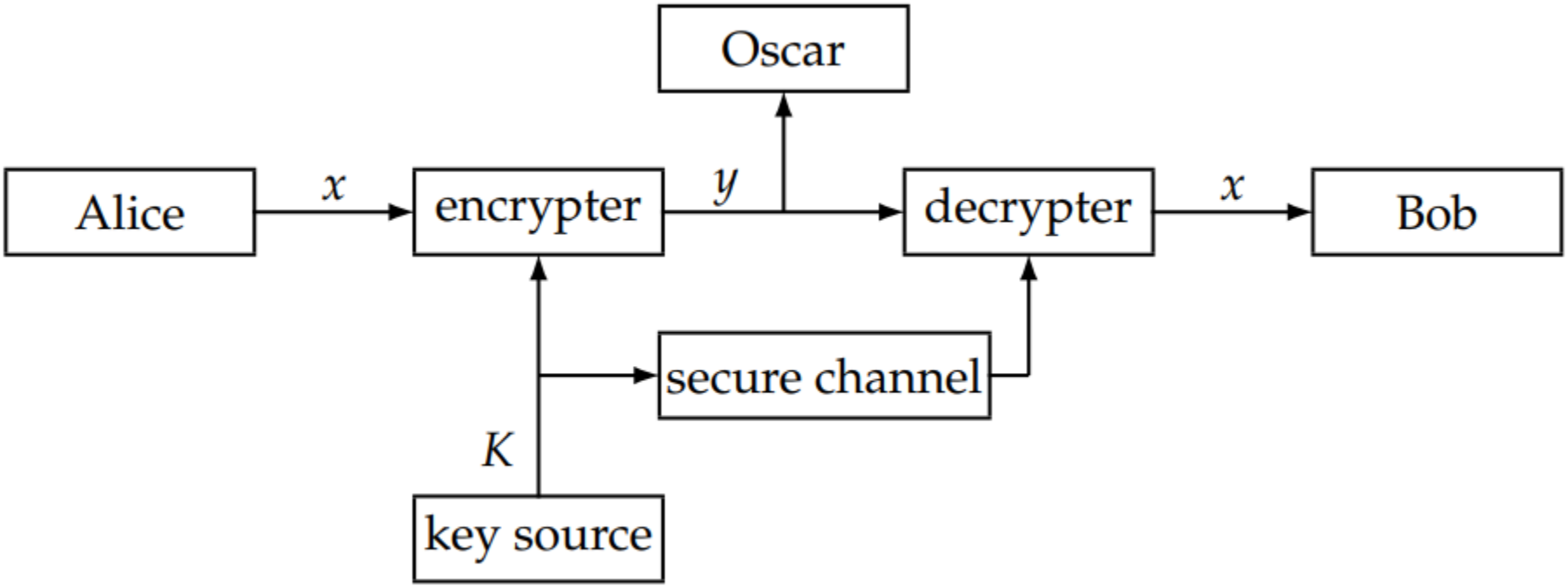
$$X = \begin{cases} a & \text{with probability } \frac{1}{2} \\ b & \text{with probability } \frac{1}{4} \\ c & \text{with probability } \frac{1}{8} \\ d & \text{with probability } \frac{1}{8} \end{cases}$$

$$H(X) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log\frac{1}{4} - \frac{1}{8}\log\frac{1}{8} - \frac{1}{8}\log\frac{1}{8} = \frac{7}{4} \text{ bits}$$

ENTROPY

- Suppose that we wish to determine the value of X with the minimum number of binary questions.
- An efficient first question is "Is $X = a$?"
- This splits the probability in half. If the answer to the first question is no, the second question can be "Is $X = b$?" The third question can be "Is $X = c$?"
- The resulting expected number of binary questions required is 1.75.
- This turns out to be the minimum expected number of binary questions required to determine the value of X in average.

- Block Cipher:
  - DES
  - AES
- Public key cryptography


Oscar
Alice
x
encrypter
y
decrypter
x
Bob
secure channel
K
key source

substitution

- A substitution cipher is a method of encrypting in which units of plaintext are replaced with the ciphertext, with the help of a key, the "units" may be single letters (the most common), pairs of letters, triplets of letters, mixtures of the above, and so forth.
- The receiver deciphers the text by performing the inverse substitution process to extract the original message.

Substitution

- The Shift Cipher is a special case of the Substitution Cipher.

- In the shift cipher, the key defines a map from each letter of the (plaintext) alphabet to some letter of the (ciphertext) alphabet, where the map is a fixed shift determined by the key

The shift cipher

- Consider encrypting English text
- Associate ‘a’ with 0; ‘b’ with 1; …; ‘z’ with 25

- $k \in \mathcal{K} = \{0, \ldots, 25\}$
- To encrypt using key k, shift every letter of the plaintext by k positions (with wraparound)
- Decry

```
helloworldz
ccccccccccc
jgnnqyqtnfb
```

Permutation

- Substitution: plaintext characters are replaced by different ciphertext characters.
- The idea of a permutation cipher is to keep the plaintext characters unchanged, but to alter their positions by rearranging them using a permutation.
- A *permutation* of a finite set X is a bijective function π: $X \rightarrow X$. In other words, the function π is one-to-one (injective) and onto (surjective). It follows that, for every $x \in X$, there is a unique element $x' \in X$ such that $\pi(x') = x$. This allows us to define the inverse permutation, π^{-1}: $X \rightarrow X$ by the rule

$$\pi^{-1}: X = X \text{ if and only if } \pi(x') = x$$

تابع یک به یک و پوشا

Permutation Cipher (also known as the Transposition Cipher)

Permutation Cipher

Let m be a positive integer. Let $P = C = (\mathbb{Z}_{26})^m$ and let K consist of all permutations of $\{1, \ldots, m\}$. For a key (i.e., a permutation) π, we define

$$e_\pi(x_1, \ldots, x_m) = (x_{\pi(1)}, \ldots, x_{\pi(m)})$$

and

$$d_\pi(y_1, \ldots, y_m) = (y_{\pi^{-1}(1)}, \ldots, y_{\pi^{-1}(m)})$$

Permutation Cipher

- Suppose m = 6 and the key is the following permutation p:

x	1	2	3	4	5	6
$\pi(x)$	3	5	1	6	4	2

.

- suppose we are given the plaintext:

```
shesellsseashellsbytheseashore.
```

Permutation Cipher

- We first partition the plaintext into groups of six letters:

shesel | lsseas | hellsb | ythese | ashore

- Now each group of six letters is rearranged according to the permutation p, yielding the following:

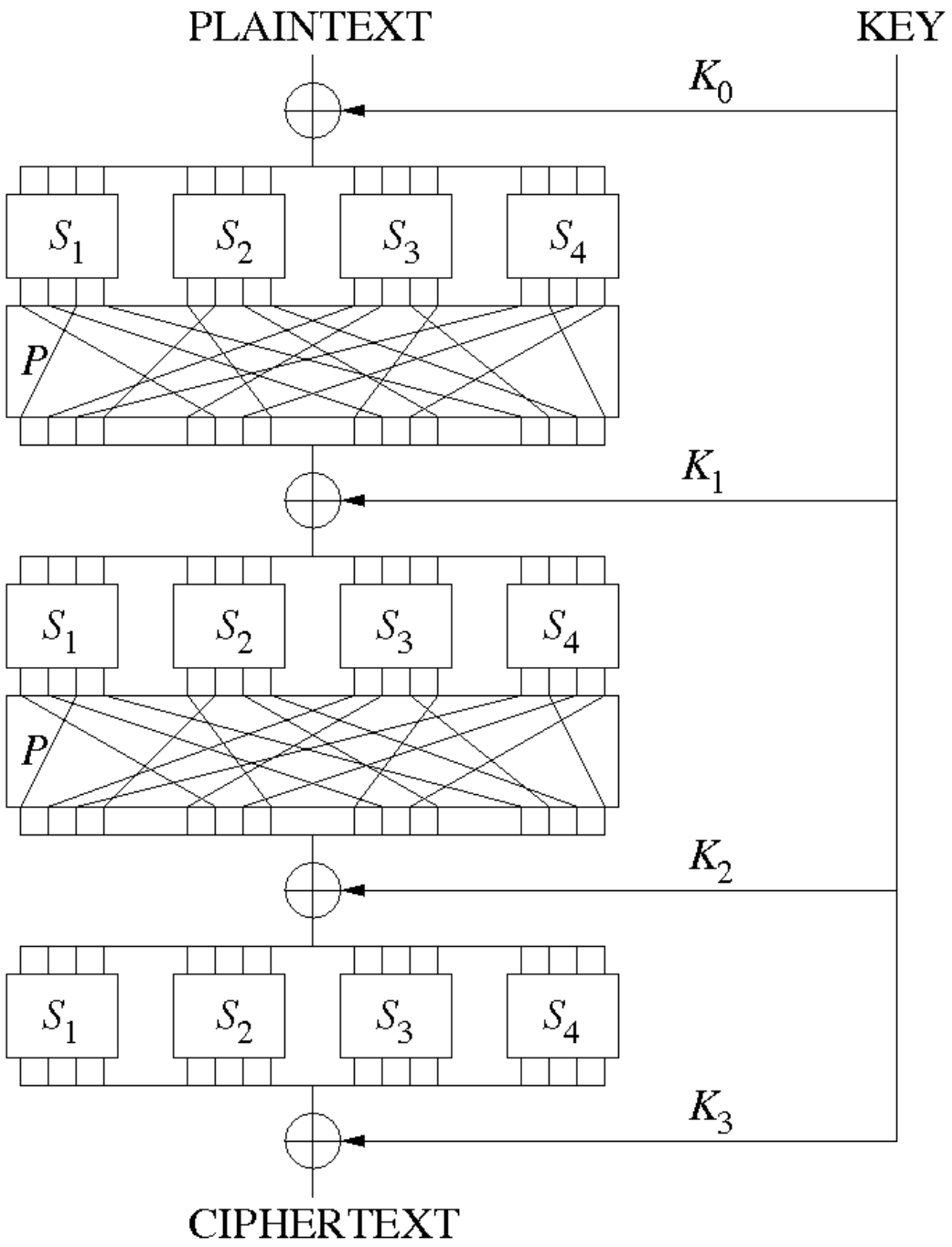
EESLSH | SALSES | LSHBLE | HSYEET | HRAEOS

x	1	2	3	4	5	6
$\pi(x)$	3	5	1	6	4	2

.

Substitution–permutation network

- An SPN-network is a series of linked mathematical operations used in block cipher algorithms such as AES (Rijndael), 3-Way, Kalyna, Kuznyechik, PRESENT, SAFER, SHARK, and Square.
- Such a network takes a block of the plaintext and the key as inputs, and applies several alternating rounds or layers of substitution boxes (S-boxes) and permutation boxes (P-boxes) to produce the ciphertext block.
- The S-boxes and P-boxes transform blocks of input bits into output bits. It is common for these transformations to be operations that are efficient to perform in hardware, such as exclusive or (XOR) and bitwise rotation. The key is introduced in each round, usually in the form of "round keys" derived from it. (In some designs, the S-boxes themselves depend on the key.)


PLAINTEXT
KEY
K_0
S_1
S_2
S_3
S_4
P
K_1
S_1
S_2
S_3
S_4
P
K_2
S_1
S_2
S_3
S_4
K_3
CIPHERTEXT

Advanced Encryption Standard

- The Advanced Encryption Standard (AES), also known by its original name Rijndael is a specification for the encryption of electronic data established by the U.S. National Institute of Standards and Technology (NIST) in 2001.
- AES is a variant of the Rijndael block cipher developed by two Belgian cryptographers, Joan Daemen and Vincent Rijmen, who submitted a proposal to NIST during the AES selection process.
- Rijndael is a family of ciphers with different key and block sizes. For AES, NIST selected three members of the Rijndael family, each with a block size of 128 bits, but three different key lengths: 128, 192 and 256 bits.

Advanced Encryption Standard

- AES has been adopted by the U.S. government. It supersedes the Data Encryption Standard (DES), which was published in 1977.
- AES ratified as a standard by National Institute of Standards and Technology of the United States (NIST), was chosen using a process lasting from 1997 to 2000 that was markedly more open and transparent than its predecessor, the DES.
- This process won praise from the open cryptographic community, and helped to increase confidence in the security of the winning algorithm from those who were suspicious of backdoors in the predecessor, DES.

Advanced Encryption Standard

Rounds one and two:

- In the nine months that followed, fifteen designs were created and submitted from several countries.
- They were, in alphabetical order: CAST-256, CRYPTON, DEAL, DFC, E2, FROG, HPC, LOKI97, MAGENTA, MARS, RC6, Rijndael, SAFER+, Serpent, and Twofish.

Advanced Encryption Standard

- The key size used for an AES cipher specifies the number of transformation rounds
  - 10 rounds for 128-bit keys.
  - 12 rounds for 192-bit keys.
  - 14 rounds for 256-bit keys.
- For AES-128, the key can be recovered with a computational complexity of $2^{126.1}$ using the biclique attack (meet-in-the-middle). For biclique attacks on AES-192 and AES-256, the computational complexities of $2^{189.7}$ and $2^{254.4}$ respectively apply.
- Related-key attacks can break AES-256 and AES-192 with complexities $2^{99.5}$ and 2^{176} in both time and data, respectively.
- Side-channel attacks

Data Encryption Standard

- Developed in the early 1970s at IBM and based on an earlier design by Horst Feistel, the algorithm was submitted to the National Bureau of Standards (NBS) following the agency's invitation to propose a candidate for the protection of sensitive, unclassified electronic government data. In 1976, after consultation with the National Security Agency (NSA), the NBS selected a slightly modified version (strengthened against differential cryptanalysis, but weakened against brute-force attacks), which was published as an official Federal Information Processing Standard (FIPS) for the United States in 1977.

Attacks faster than brute force on DES

There are three attacks known that can break the full 16 rounds of DES with less complexity than a brute-force search:

- Differential cryptanalysis
- Linear cryptanalysis
- Davies' attack

- The attacks are theoretical and are generally considered infeasible to mount in practice.
- Differential cryptanalysis was rediscovered in the late 1980s by Eli Biham and Adi Shamir;
- it was known earlier to both IBM and the NSA and kept secret. To break the full 16 rounds, differential cryptanalysis requires 2^47 chosen plaintexts.

Key Differences Between DES and AES

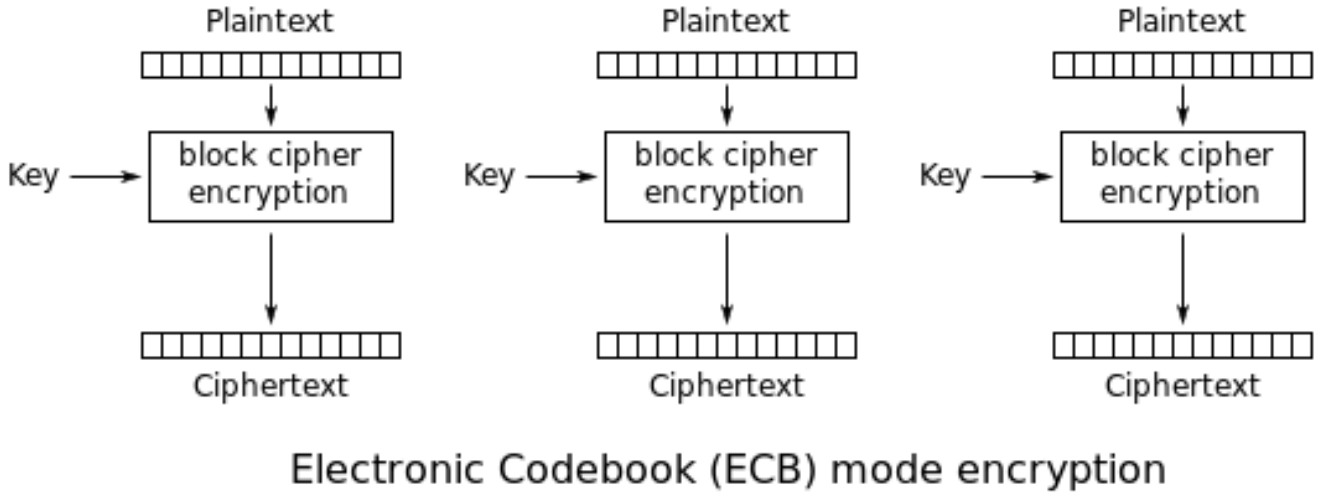
- The main difference between DES and AES is that in DES, the block is split into two halves before being processed further, but in AES, the entire block is processed to get ciphertext.
- DES has a key size of 56 bits, which is less than AES, which has a secret key size of 128, 192, or 256
- AES is comparatively faster than DES.
- The smaller key size of DES makes it less secure than AES.
- The Feistel Cipher principle is used in the DES algorithm, while the substitution and permutation principle are used in the AES

Block cipher mode of operation

- Block cipher mode of operation is an algorithm that uses a block cipher to provide information security such as confidentiality or authenticity.
- A block cipher by itself is only suitable for the secure cryptographic transformation (encryption or decryption) of one fixed-length group of bits called a block.
- A mode of operation describes how to repeatedly apply a cipher's single-block operation to securely transform amounts of data larger than a block

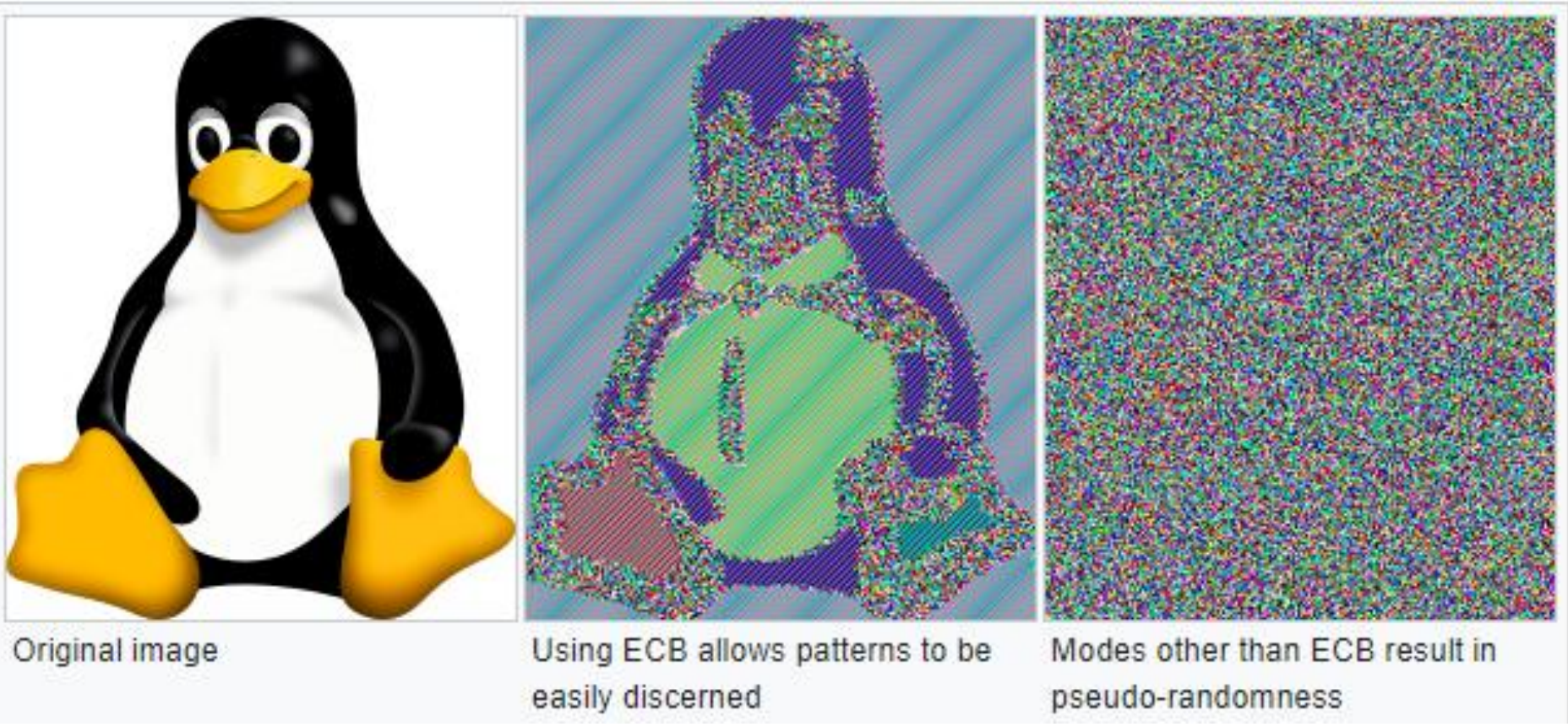
Block cipher mode of operation

- Electronic Codebook (ECB):



Electronic Codebook (ECB) mode encryption

Block cipher mode of operation



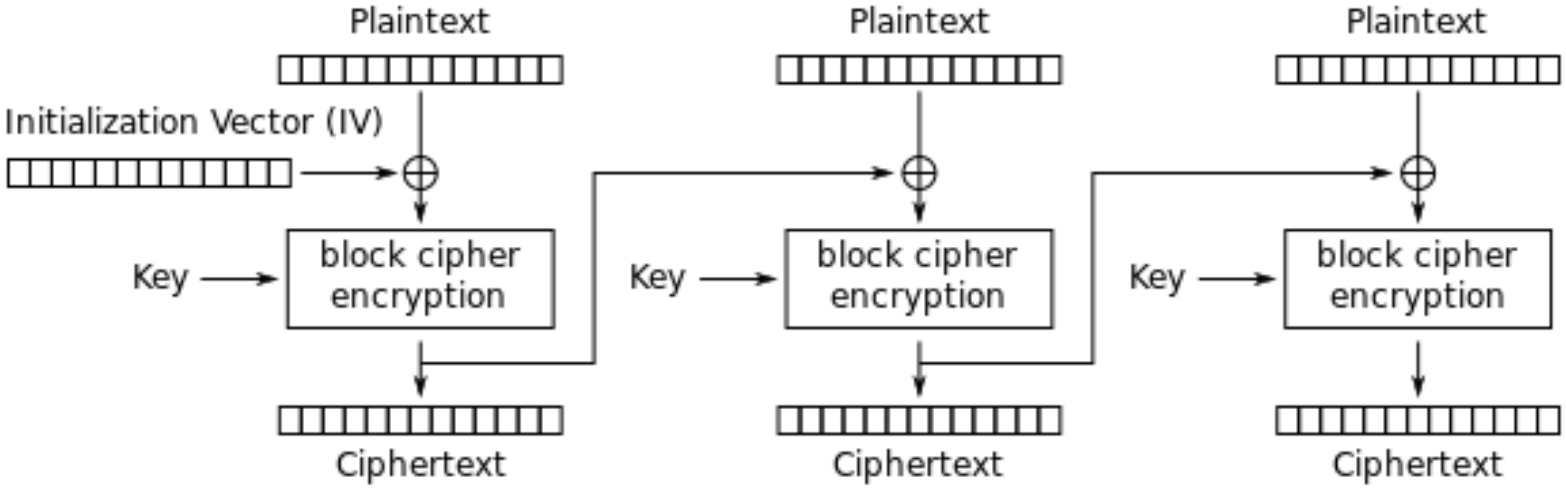
Original image

Using ECB allows patterns to be easily discerned

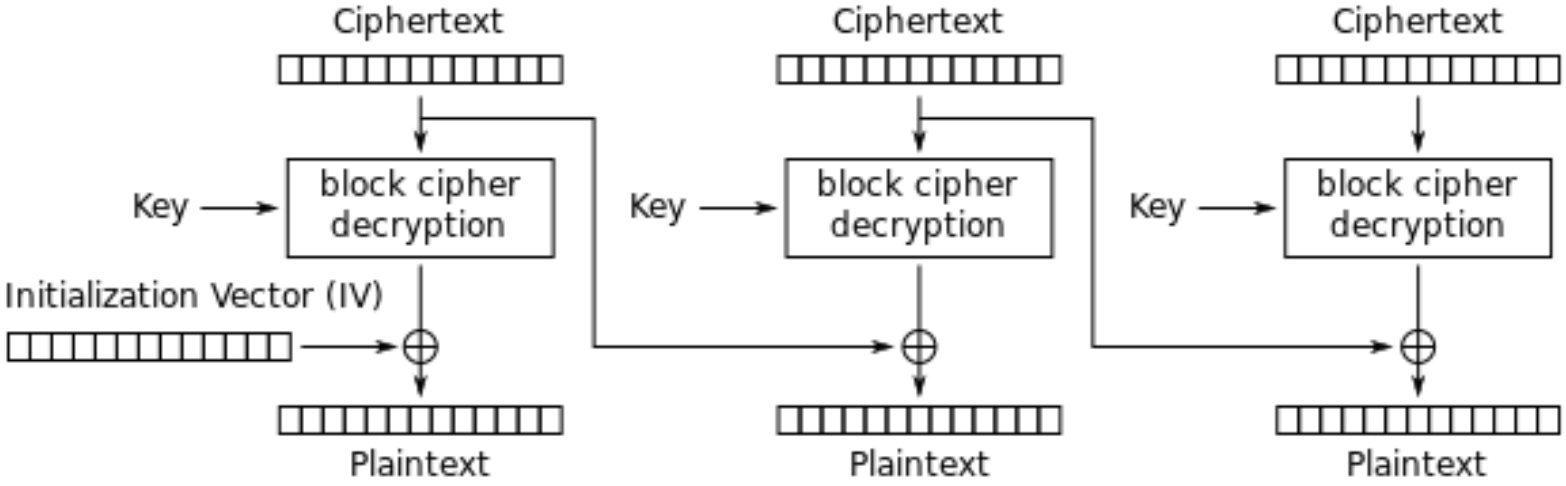
Modes other than ECB result in pseudo-randomness

Block cipher mode of operation

Cipher block Chaining (CBC)

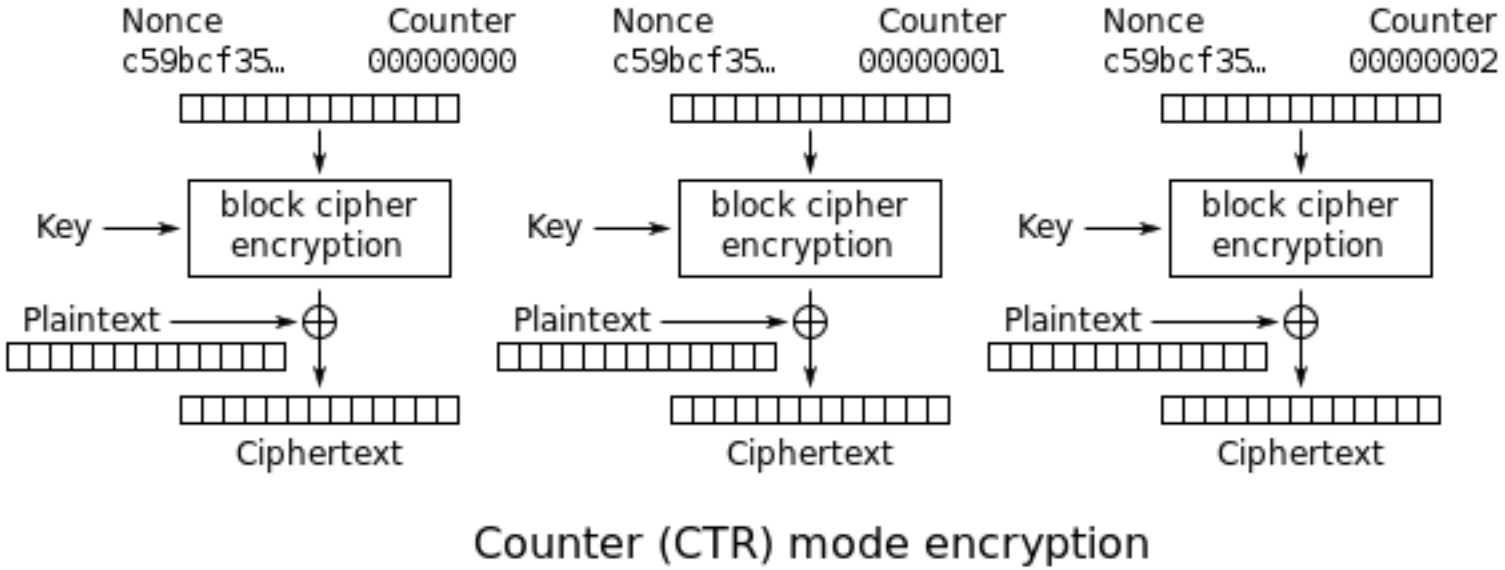


Cipher Block Chaining (CBC) mode encryption

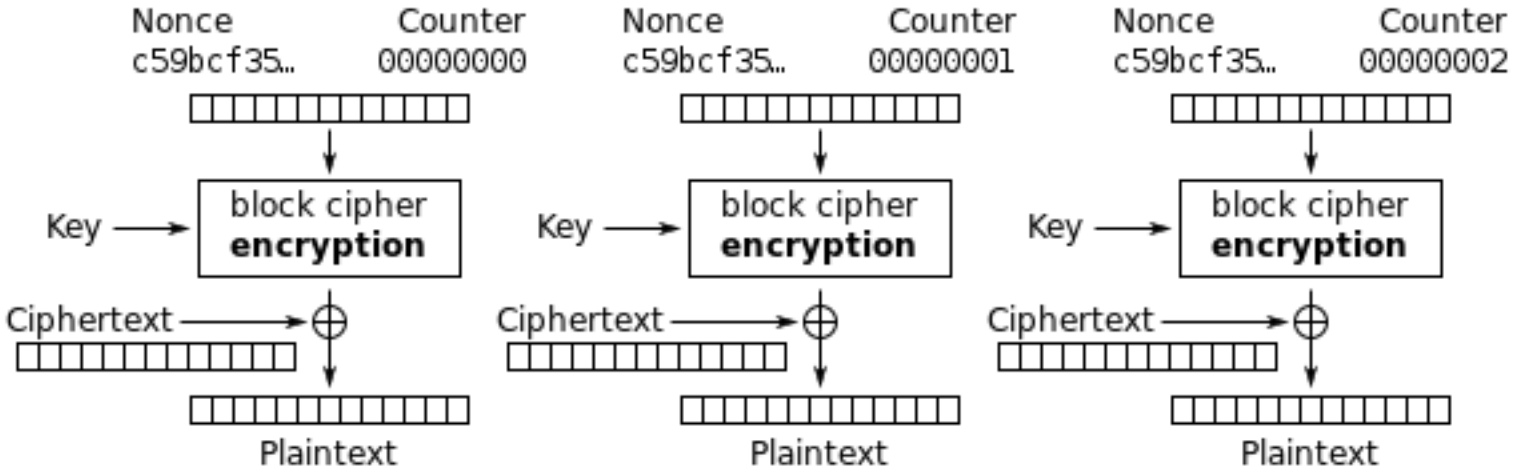


Block cipher mode of operation

Counter (CTR)



Counter (CTR) mode encryption

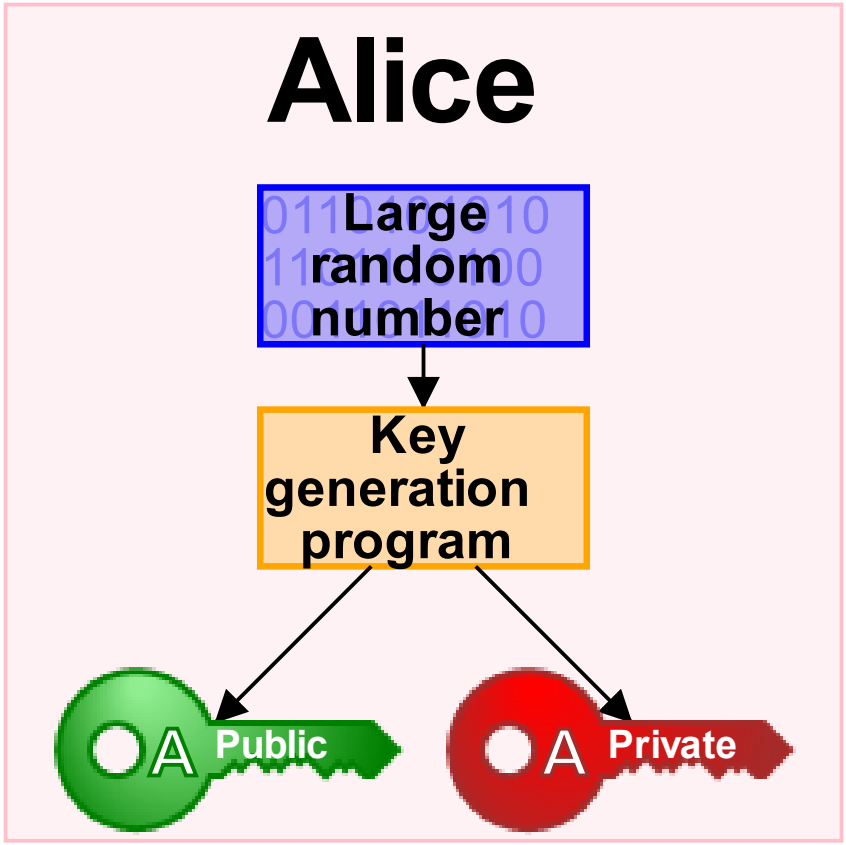


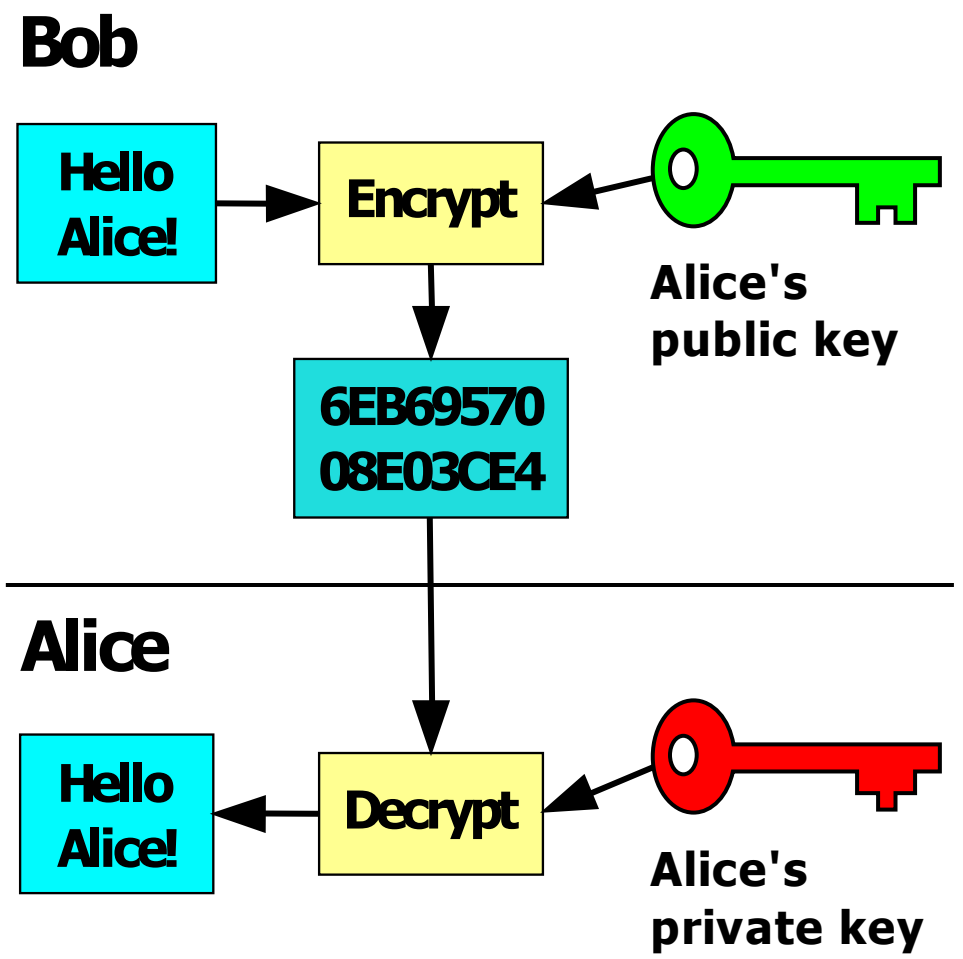
Counter (CTR) mode decryption

Public-key cryptography

- Asymmetric cryptography, is the field of cryptographic systems that use pairs of related keys. Each key pair consists of a public key and a corresponding private key
- Security of public-key cryptography depends on keeping the private key secret; the public key can be openly distributed without compromising security
- In a public-key encryption system, anyone with a public key can encrypt a message, yielding a ciphertext, but only those who know the corresponding private key can decrypt the ciphertext to obtain the original message

Public-key cryptography





Digital signature