Overview of Cryptography

Prof: Pierre-Alain Fouque

Reference: http://cacr.uwaterloo.ca/hac/about/chap1.pdf

Vernam Ciphers a.k.a. One-Time Pad

- Let A={0,1}. A binary message $m_1m_2 \ldots m_t$ is operated by a binary key $k_1k_2 \ldots k_t$ of the same length to produce the ciphertext string $c_1c_2 \ldots c_t$:
- $c_i = m_i \oplus k_i$ for all i=1…t.
- To decrypt, it suffices to compute $c_i \oplus k_i$ to recover m_i.

- If we encrypt two messages with the same key, we can obtain

$c \oplus c' = m \oplus m'$, where $c = m \oplus k$ and $c' = m' \oplus k$ since $k \oplus k = 0$ and $k \oplus 0 = k$.

Unbreakable cipher used during the cold war: unconditional security

Main drawbacks:

1. the key must be random,
2. key must be as long as the message, and
3. key must be changed for each message.

Unconditional vs. Computational Security

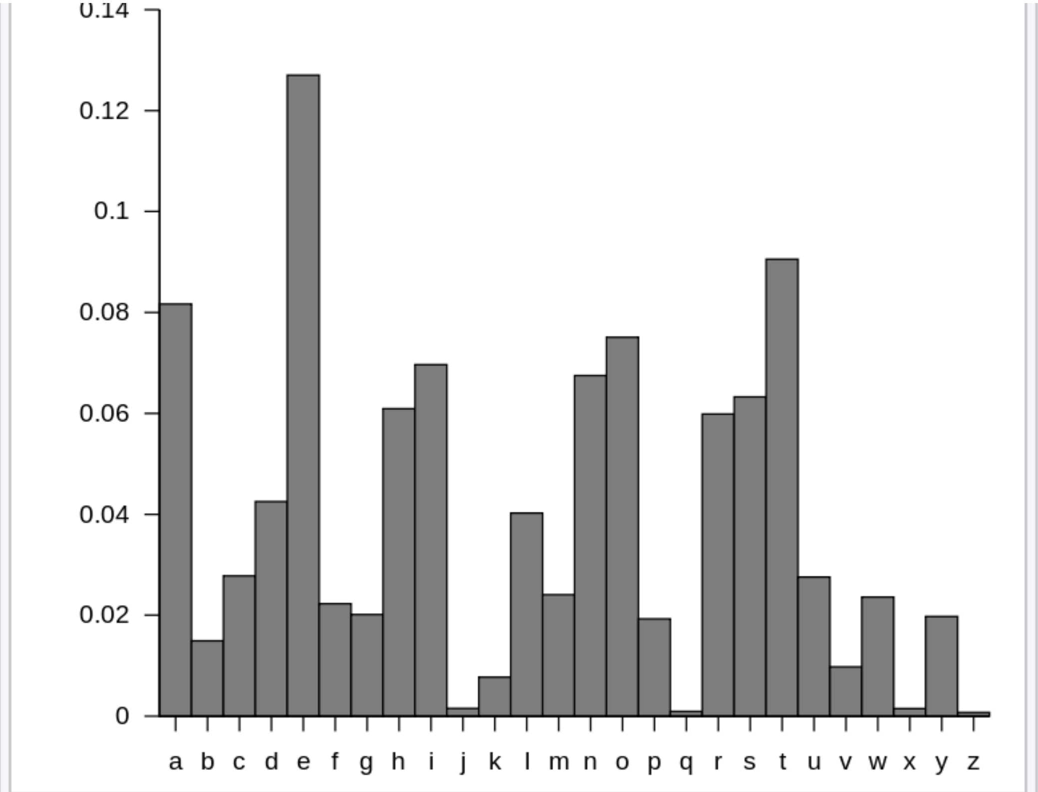
- A powerful adversary with infinite time cannot obtain information about the plaintext given only the ciphertext
- If a ciphertext c is obtained, anyone can produce a plaintext p and a key k such that $c=p\oplus k$ for any plaintext $|p|=|c|$
- The adversary cannot distinguish plaintext with equal length

- However, in practice keys are reused across several ciphertexts

- In a practical point of view, computational security is preferred: it is computationally hard to recover the plaintext (but possible for an adversary with infinite time…).
- E.g.: for a block cipher, we can exhaust all keys 2^{128} operations

Vigenere Cipher

- Vigenere is based on the same idea as Vernam with {a,b,c,…z} alphabet or ascii characters



Les fréquences relatives des lettres dans le texte

- Vigenere cipher reuses the same key
- IC (index of coincidence) = $\sum_{a \in A} p_a$ where

p_a is the probability of character a

- IC is invariant with substitution
- IC is higher when the distribution is not

uniform

Cryptanalysis Vigenere Cipher

- Assume the key length is known, one can extract substrings encrypted with the same letter
- Such encryption is called a shift encryption since the whole alphabet is shifted
- Easier to break than substitution: once the encryption of one letter is known, we can deduce all the substitution

- Learning the length: Guess all length and compute the IC
- ICM: $\sum_{a \in A} p_a p'_a$ where p_a and p'_a are the probabilities of two strings

The key space

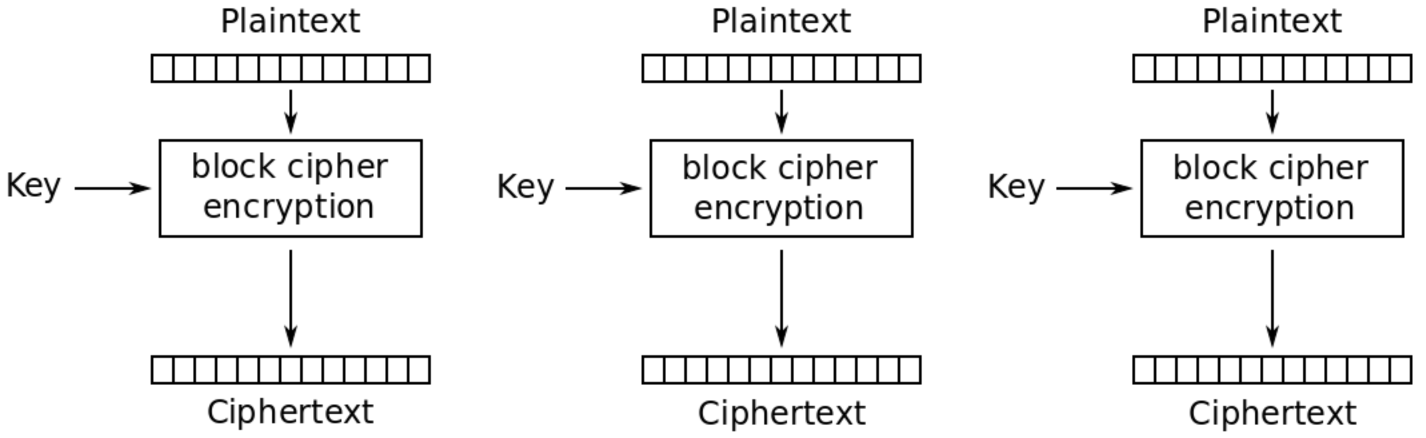
- The size of the key space is the number of encryption/decryption key pairs available in the cipher system. A key is a compact way to specify the encryption function (from the set of all encryption functions).

- E.g. a substitution of block length t has $(2^t)!$ encryption functions

- A necessary, but usually not sufficient, condition for an encryption scheme to be secure is that the key space be large enough to preclude exhaustive search. E.g. $26! \approx 4 \times 10^{26}$.

Encrypting long messages

- Mode of operations: how using a block cipher to encrypt large messages ?



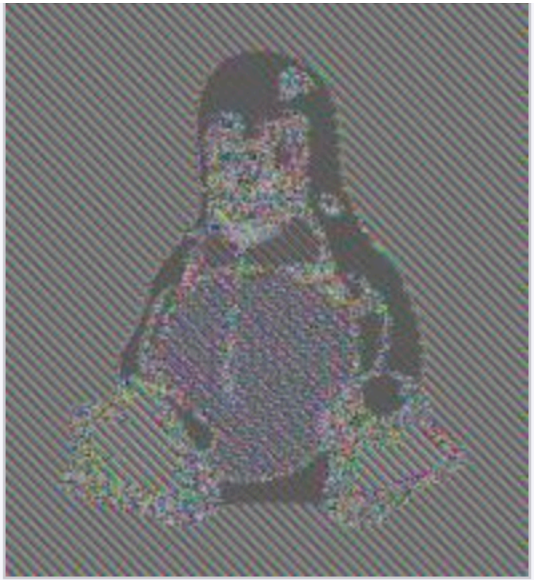
Electronic Codebook (ECB) mode encryption

Problems:
- deterministic

Other modes: CFB, OFB, PCBC, CTS (Ciphertext stealing)

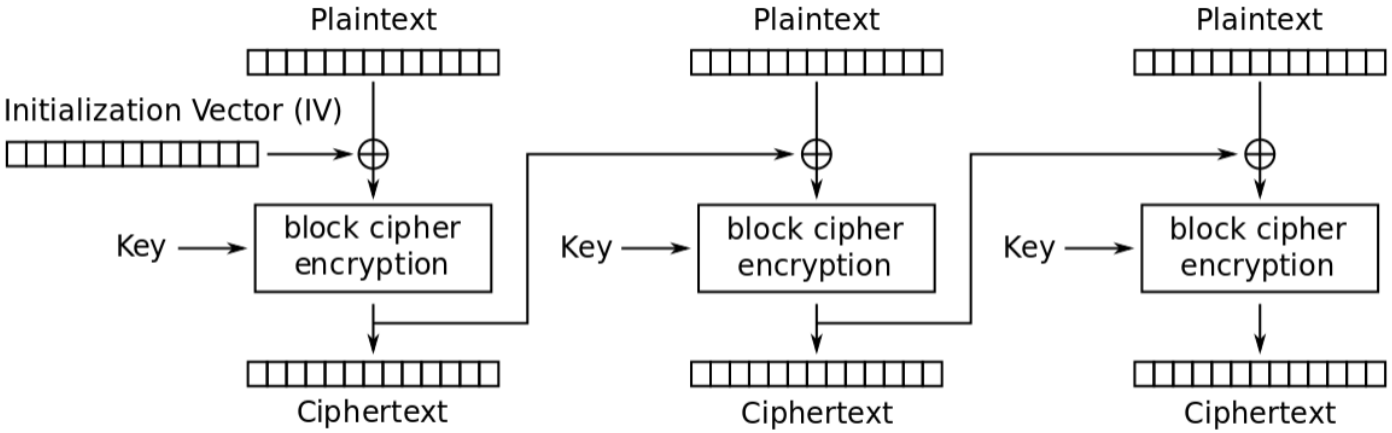
Electronic Code Book is deterministic … CBC better

Original image



Encrypted using ECB mode

Randomization is useful …



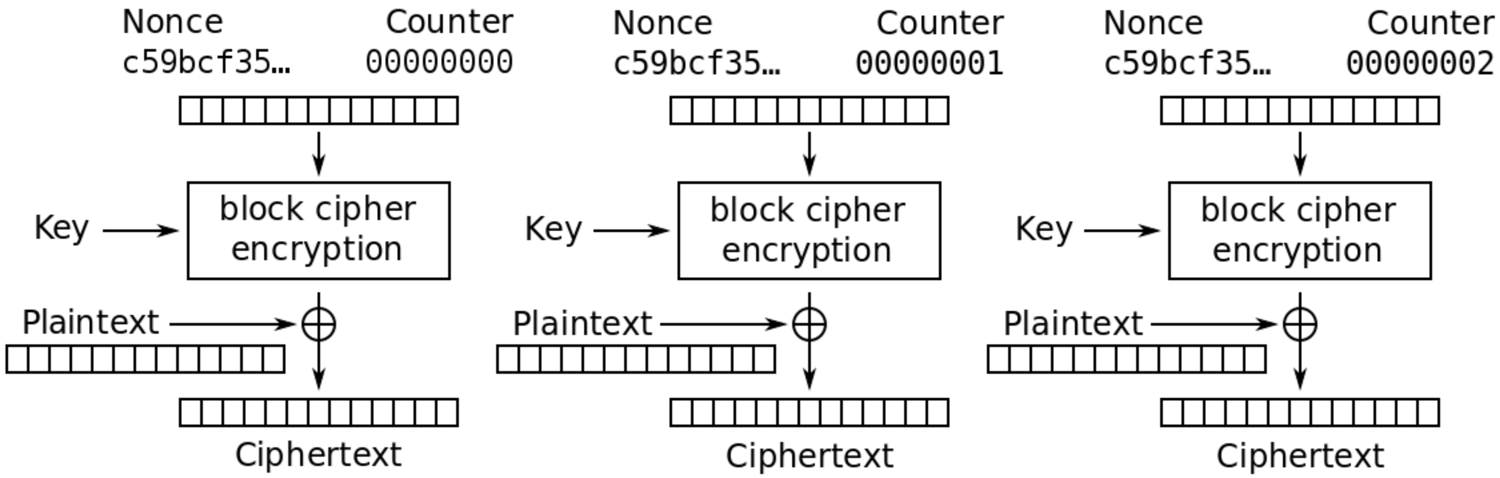
Cipher Block Chaining (CBC) mode encryption

Stream Cipher and CTR mode of operations

Vernam cipher is unconditionally secure

Main problem: key reuse

- Generate the key with a smaller one using a pseudorandom number generator : output looks random but bitstring generated deterministically with from a secrete seed
- Cryptographically secure pseudorandom generator are hard to design: rand from c language is not good
- Block cipher can be used as follows
- If stream cipher are resymchronized, same key is generated (WPA)
- In order to make it stateless, a nonce is usually added to generate different keystream



Counter (CTR) mode encryption